

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

3M nucleic - nucleic search, using sw model

run on: March 11, 2004, 03:37:21 ; Search time 353.257 Seconds  
(without alignments)  
15799.545 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgttttggacatagct.....gctgaataataatgncgtc 1516

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	99.9	1516	9	US-09-919-172-38
2	1515	99.9	1516	9	US-09-974-298-92
3	1515	99.9	1516	14	US-10-084-817-8
4	1389.4	91.6	1399	9	US-09-864-864-321
5	1387.8	91.5	1399	9	US-09-954-456-2119
6	1387.8	91.5	1399	14	US-10-240-965-2
7	1387.8	91.5	1399	14	US-10-269-909-2
8	1387.8	91.5	1399	15	US-10-133-337-2
9	1387.8	91.5	1399	15	US-10-159-563-2
10	1379.4	91.0	1431	9	US-09-925-301-46
11	1367.4	90.2	1377	14	US-10-210-120-18
12	886.8	58.5	1402	9	US-09-917-800A-1585
13	886.8	58.5	1402	14	US-10-316-253-299
14	822	58.2	2500	10	US-08-814-353-20447
15	597.4	39.4	600	12	US-10-085-783A-51438

16	597.4	39.4	600	15	US-10-242-535A-51438	Sequence 51438, A
17	589.4	38.9	614	9	US-09-864-864-241	Sequence 211, App
18	583.4	38.5	618	12	US-10-085-783A-13898	Sequence 13898, A
19	583.4	38.5	618	15	US-10-242-535A-13898	Sequence 13898, A
20	529	34.9	592	12	US-10-085-783A-48352	Sequence 48352, A
21	529	34.9	592	15	US-10-242-535A-48352	Sequence 48352, A
22	519.4	34.3	537	9	US-09-777-564-134	Sequence 134, App
23	519.4	34.3	537	14	US-10-015-219-134	Sequence 134, App
24	505	33.3	557	12	US-10-085-783A-48313	Sequence 48313, A
25	505	33.3	557	15	US-10-242-535A-48313	Sequence 48313, A
26	500	33.0	530	12	US-10-085-783A-54108	Sequence 54108, A
27	500	33.0	530	15	US-10-242-535A-54108	Sequence 54108, A
28	496	32.7	590	12	US-10-085-783A-48227	Sequence 48227, A
29	496	32.7	590	15	US-10-242-535A-48227	Sequence 48227, A
30	491.4	32.4	500	12	US-10-085-783A-23999	Sequence 23999, A
31	491.4	32.4	500	15	US-10-242-535A-23999	Sequence 23999, A
32	482	31.8	495	12	US-10-085-783A-38267	Sequence 38267, A
33	482	31.8	495	15	US-10-242-535A-38267	Sequence 38267, A
34	474	31.3	499	12	US-10-085-783A-26128	Sequence 26128, A
35	474	31.3	499	15	US-10-242-535A-26128	Sequence 26128, A
36	468.2	30.9	473	12	US-10-085-783A-24370	Sequence 24370, A
37	468.2	30.9	473	15	US-10-242-535A-24370	Sequence 24370, A
38	460.2	30.4	470	12	US-10-085-783A-42942	Sequence 42942, A
39	460.2	30.4	470	15	US-10-242-535A-42942	Sequence 42942, A
40	459.4	30.3	461	9	US-09-998-598-607	Sequence 607, App
41	458	30.2	473	12	US-10-085-783A-31104	Sequence 31104, A
42	458	30.2	473	15	US-10-242-535A-31104	Sequence 31104, A
43	457.8	30.2	461	9	US-09-998-598-1129	Sequence 1129, App
44	455.2	30.0	472	12	US-10-085-783A-55749	Sequence 55749, A
45	455.2	30.0	472	15	US-10-242-535A-55749	Sequence 55749, A

## ALIGNMENTS

### RESULT 1

US-09-919-172-38  
; Sequence 38, Application US/09919172  
; Patent No. US20020119463A1  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Mary  
; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
; FILE REFERENCE: PA-0036 US  
; CURRENT APPLICATION NUMBER: US/09/919,172  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/222,469  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PERL Program  
; SEQ ID NO 38  
; LENGTH: 1516  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1303785CBI  
; NAME/KEY: unsure  
; LOCATION: 1512  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-919-172-38

Query Match 99.9%; Score 1515; DB 9; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGCAATTAATCT 60  
Db 1 CTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGCAATTAATCT 60  
QY 61 TCTGTTGCTAGTGCTTCTTTAAATCTTAAATCAGAGCCCAAGTCTCCAC 120  
Db 61 TCTGTTGCTAGTGCTTCTTTAAATCTTAAATCAGAGCCCAAGTCTCCAC 120





121 TGCAGTGTGAAATCTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 180  
 121 TGCAGTGTGAAATCTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 180  
 181 AAGAGACATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 240  
 181 AAGAGACATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 240  
 241 GAAATGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 300  
 241 GAAATGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 300  
 301 GCGGTGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 360  
 301 GCGGTGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 360  
 361 AATATGTTTAAAGGTGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 420  
 361 AATATGTTTAAAGGTGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 420  
 421 GCAAGAGTCAAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 480  
 421 GCAAGAGTCAAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 480  
 481 ACATGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 540  
 481 ACATGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 540  
 541 CCAGGCAATTTGATGCTGATGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 600  
 541 CCAGGCAATTTGATGCTGATGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 600  
 601 GATATCTTAATGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 660  
 601 GATATCTTAATGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 660  
 661 GTCTACAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 720  
 661 GTCTACAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 720  
 721 GATTTTCGGAACGCTTTGCTTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 780  
 721 GATTTTCGGAACGCTTTGCTTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 780  
 781 AATGAGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 840  
 781 AATGAGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 840  
 841 GGGACAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 900  
 841 GGGACAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 900  
 901 AGAGTGTTCAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 960  
 901 AGAGTGTTCAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 960  
 961 GAGTTGAAAGGTGACATTTGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1020  
 961 GAGTTGAAAGGTGACATTTGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1020  
 1021 CCAGTGTTCAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1080  
 1021 CCAGTGTTCAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1080  
 1081 GCATTTGATGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1140  
 1081 GCATTTGATGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1140  
 1141 TATCAGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1200  
 1141 TATCAGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 1200

Appendix 3 1/2

QY 1201 TATGAGAAATCTCTGCTGCTCTTTTGTGAGGAACTAAACATTTCCCTTGTATGCTCTCAA 1260  
 DB 1201 TATGAGAAATCTCTGCTGCTCTTTTGTGAGGAACTAAACATTTCCCTTGTATGCTCTCAA 1260  
 QY 1261 GCTATGATCAGAGAGATTTTAAATATATATTTTTCATCTTAAAGCTTAAATAGGAAAGTTT 1320  
 DB 1261 GCTATGATCAGAGAGATTTTAAATATATATTTTTCATCTTAAAGCTTAAATAGGAAAGTTT 1320  
 QY 1321 CTTCAACAGAGATTTACAGTGTAGTCTACCTACATCTGCTGAAATAATATAGCTTTAAATCATTT 1380  
 DB 1321 CTTCAACAGAGATTTACAGTGTAGTCTACCTACATCTGCTGAAATAATATAGCTTTAAATCATTT 1380  
 QY 1381 TTATATTATAAATCTGTATATAATAGAGATAAGTCCATTTTAAATAATGTTTTTCCCAAAAC 1440  
 DB 1381 TTATATTATAAATCTGTATATAATAGAGATAAGTCCATTTTAAATAATGTTTTTCCCAAAAC 1440  
 QY 1441 CATAAACCCCTATACAAAGTTGTTCTTAGTAAACAATATAGAGAAAGATGCTATGTAGCTG 1500  
 DB 1441 CATAAACCCCTATACAAAGTTGTTCTTAGTAAACAATATAGAGAAAGATGCTATGTAGCTG 1500  
 QY 1501 AAAATAAATGNCGTC 1516  
 DB 1501 AAAATAAATGNCGTC 1516

RESULT 2  
 US-09-974-298-92  
 ; Sequence 92, Application US/09974298  
 ; Patent No. US20020156263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, Huel-Mei  
 ; TITLE OF INVENTION: GENES EXPRESSED IN BLAST CANCER  
 ; FILE REFERENCE: PA-0037 P  
 ; CURRENT APPLICATION NUMBER: US-09/974,298  
 ; CURRENT FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: 60/238,331  
 ; PRIOR FILING DATE: 2000-05-10  
 ; NUMBER OF SEQ ID NOS: 194  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 92  
 ; LENGTH: 1516  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 1303785CB1  
 ; NAME/KEY: unsure  
 ; LOCATION: 1512  
 ; OTHER INFORMATION: a, t, c, g, or other  
 ; US-09-974-298-92

Query Match 99.9%; Score 1515; DB 9; Length 1516;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGTCTTGTAGTGTGCTTTCTTTAAATCTTATATAATCAGAGCCCAAGTCTCCAC 120  
 DB 1 CTTTGTCTTGTAGTGTGCTTTCTTTAAATCTTATATAATCAGAGCCCAAGTCTCCAC 120  
 QY 61 TCTGTCTCTAGTGTGCTTTCTTTAAATCTTATATAATCAGAGCCCAAGTCTCCAC 120  
 DB 61 TCTGTCTCTAGTGTGCTTTCTTTAAATCTTATATAATCAGAGCCCAAGTCTCCAC 120  
 QY 121 TGCAGTGTGAAATCTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 180  
 DB 121 TGCAGTGTGAAATCTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 180  
 QY 181 AAGAGACATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 240  
 DB 181 AAGAGACATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 240  
 QY 241 GAAATGAGAGAGAGAGATTTTTCAGAGAAATTTCTCTTTAGTTCTTTTTCAGAGAGGTAGAT 300

b	241	GAAAA	TGAAG	CGACG	GAATAT	GTTC	CAAA	CTGT	GAAGT	CATC	CAAA	GGTGTG	GTCC	CGGA	TCA	300															
y	301	GC	GGT	GAG	CC	CTAT	CC	TAC	CTT	CAAT	CTC	TGG	ATG	CG	CTCT	GC	CA	TAA	GGCC	360											
b	301	GC	GGT	GAG	CC	CTAT	CC	TAC	CTT	CAAT	CTC	TGG	ATG	CG	CTCT	GC	CA	TAA	GGCC	360											
y	361	ATA	T	GGT	TAA	AGG	T	CG	AT	GAAG	CA	CC	AT	CA	TTT	CT	AA	TG	AG	GA	CA	AT	420								
b	361	ATA	T	GGT	TAA	AGG	T	CG	AT	GAAG	CA	CC	AT	CA	TTT	CT	AA	TG	AG	GA	CA	AT	420								
y	421	GCA	CAG	CGT	CAA	CAG	AT	CA	AA	GAG	CA	TAT	CT	CAC	GGA	AA	CA	AG	AA	GG	CC	CT	GC	AT	GA	AA	480				
b	421	GCA	CAG	CGT	CAA	CAG	AT	CA	AA	GAG	CA	TAT	CT	CAC	GGA	AA	CA	AG	AA	GG	CC	CT	GC	AT	GA	AA	480				
y	481	AC	ACT	G	AA	GA	AG	CC	CTT	A	CAG	GT	CAC	TTG	GAG	AG	TT	GT	TT	AG	CT	CT	G	T	CT	CT	AA	AA	CT	540	
b	481	AC	ACT	G	AA	GA	AG	CC	CTT	A	CAG	GT	CAC	TTG	GAG	AG	TT	GT	TT	AG	CT	CT	G	T	CT	CT	AA	AA	CT	540	
y	541	CC	AG	CG	CA	ATT	GA	T	GT	GA	T	GA	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	600	
b	541	CC	AG	CG	CA	ATT	GA	T	GT	GA	T	GA	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	600	
y	601	GAT	ACT	CT	TA	ATT	G	AG	ATTT	T	GG	CA	CT	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	660
b	601	GAT	ACT	CT	TA	ATT	G	AG	ATTT	T	GG	CA	CT	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	AA	CT	TA	CA	AG	660
y	661	GT	CTA	CAG	AG	GA	AA	CT	GA	AG	AT	CT	GG	CA	AG	CA	TAT	AA	CC	CT	CAG	AC	CA	CA	CA	CA	CA	CA	CA	CA	720
b	661	GT	CTA	CAG	AG	GA	AA	CT	GA	AG	AT	CT	GG	CA	AG	CA	TAT	AA	CC	CT	CAG	AC	CA	CA	CA	CA	CA	CA	CA	CA	720
y	721	GATTTT	CG	AA	CG	CTT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	780	
b	721	GATTTT	CG	AA	CG	CTT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	780	
y	781	AAT	GA	AG	ACT	TT	GG	CT	GA	TT	TC	AG	AT	CC	CA	AG	CA	TG	AT	GA	AG	CA	GAG	AA	GG	AA	GG	AA	GG	AA	840
b	781	AAT	GA	AG	ACT	TT	GG	CT	GA	TT	TC	AG	AT	CC	CA	AG	CA	TG	AT	GA	AG	CA	GAG	AA	GG	AA	GG	AA	GG	AA	840
y	841	GGG	CAG	AG	CT	TA	AG	CG	TT	CA	AT	ACC	AT	CT	TAC	CA	AG	AA	CG	TA	TCC	CA	AA	CT	TC	GC	900				
b	841	GGG	CAG	AG	CT	TA	AG	CG	TT	CA	AT	ACC	AT	CT	TAC	CA	AG	AA	CG	TA	TCC	CA	AA	CT	TC	GC	900				
y	901	AGA	GT	GT	TT	CAG	AA	T	PA	C	CA	AG	T	AG	CA	TG	AT	GA	AG	CA	TG	AT	GA	AG	CA	TG	AT	GA	AG	CA	960
b	901	AGA	GT	GT	TT	CAG	AA	T	PA	C	CA	AG	T																		

APPENDIX B  $\frac{1}{2}$

```

1381 TTATATTATAACTCTGTATAATAGACATAAGTCCATTTTAAAAATGTTTTCCCAAC 1440
1381 TTATATTATAACTCTGTATAATAGACATAAGTCCATTTTAAAAATGTTTTCCCAAC 1440
1441 CATAAAAACCCCTATACAAAGTTGTTCTAGTACAAATACATGAGAAGATGTCATGTAGCTG 1500
1441 CATAAAAACCCCTATACAAAGTTGTTCTAGTACAAATACATGAGAAGATGTCATGTAGCTG 1500
1501 AAAATAAAAAATGNCGTC 1516
1501 AAAATAAAAAATGNCGTC 1516
1501 AAAATAAAAAATGNCGTC 1516
1501 AAAATAAAAAATGNCGTC 1516
RESULT 3
US-10-084-817-8
; Sequence 8, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1303785CB1
; NAME/KEY: unsure
; LOCATION: 1512
; OTHER INFORMATION: a, t, c, g, or other
US-10-084-817-8
Query Match 99.9%; Score 1515; DB 14; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTGTGTTTTGGACATAGCTGAGCCATGTACTTCAACAGAGCGGCCCAATTACTTAAC 60
Db 1 CTTTGTGTTTTGGACATAGCTGAGCCATGTACTTCAACAGAGCGGCCCAATTACTTAAC 60
QY 61 TCTGGTGTCTAGGTTGCGCTTCCTTTAAAAATCCTATAAAAAATCAGAACCCCAAGTCTCCAC 120
Db 61 TCTGGTGTCTAGGTTGCGCTTCCTTTAAAAATCCTATAAAAAATCAGAACCCCAAGTCTCCAC 120
QY 121 TGCCAGTGTGAAATCTTCAGAGAGAATTTCTTTAGTTCTTTTGCAGAGAGGTAGAGAT 180
Db 121 TGCCAGTGTGAAATCTTCAGAGAGAATTTCTTTAGTTCTTTTGCAGAGAGGTAGAGAT 180
QY 181 AAAGACACTTTTTCAAAATTCGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTGTTTATT 240
Db 181 AAAGACACTTTTTCAAAATTCGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTGTTTATT 240
QY 241 GAAAAATGAAGAGCAGCAATATGTTCCAACTGTAAGTCATCCAAAGGTTGCCCGATCA 300
Db 241 GAAAAATGAAGAGCAGCAATATGTTCCAACTGTAAGTCATCCAAAGGTTGCCCGATCA 300
QY 301 GCGGTGAGCCCTATCCTACCTTCAATCCATCCTCGGATGTGCTGCTTGCATTAAGGCC 360
Db 301 GCGGTGAGCCCTATCCTACCTTCAATCCATCCTCGGATGTGCTGCTTGCATTAAGGCC 360
QY 361 ATAATGTTTAAAGTGTGGATGAAGAACCATCATTGACATTTCTAACTTAAGGGAACAAAT 420
Db 361 ATAATGTTTAAAGTGTGGATGAAGAACCATCATTGACATTTCTAACTTAAGGGAACAAAT 420

```

Result No.	Score	Query			DB	ID	Description
		Match	Length	\$			
1	3693	100.0	3693	4	US-09-976-594-908	Sequence 908, App	
2	892.8	24.2	826	4	US-09-621-976-15155	Sequence 15155, A	
3	619.4	16.8	621	4	US-09-643-597-209	Sequence 209, App	
4	619.4	16.8	621	4	US-09-480-884A-209	Sequence 209, App	
5	619.4	16.8	621	4	US-09-542-615A-209	Sequence 209, App	
6	619.4	16.8	621	4	US-09-606-421B-209	Sequence 209, App	
7	477.2	12.9	708	3	US-08-896-164-34	Sequence 34, Appl	
8	336	9.1	336	4	US-09-683-597-315	Sequence 315, App	
9	336	9.1	336	4	US-09-480-884A-315	Sequence 315, App	
10	336	9.1	336	4	US-09-542-615A-315	Sequence 315, App	
11	336	9.1	336	4	US-09-606-421B-315	Sequence 315, App	
12	70.2	1.9	1255	4	US-09-412-554A-7	Sequence 7, Appl	
13	70.2	1.9	3192	4	US-09-412-554A-1	Sequence 1, Appl	
14	66.8	1.8	2949	4	US-09-412-554A-3	Sequence 3, Appl	
15	53.2	1.4	7218	1	US-08-232-463-14	Sequence 14, Appl	
16	50	1.4	7218	1	US-08-232-463-14	Sequence 14, Appl	
17	47	1.3	1771	4	US-09-907-794A-158	Sequence 158, App	
18	47	1.3	1771	4	US-08-866-028-36	Sequence 36, Appl	
19	47	1.3	1771	4	US-09-905-125A-158	Sequence 158, App	
20	47	1.3	1771	4	US-09-902-775A-158	Sequence 158, App	
21	45.8	1.2	11050	4	US-10-204-708-86	Sequence 86, Appl	
22	45.6	1.2	5852	1	US-07-867-106-2	Sequence 2, Appl	
23	45	1.2	6243	2	US-09-056-075-1	Sequence 1, Appl	
24	44.2	1.2	6801	4	US-10-204-708-62	Sequence 62, Appl	
25	44.2	1.2	11049	4	US-10-204-708-22	Sequence 22, Appl	
26	43.8	1.2	1813	3	US-09-071-324-3	Sequence 3, Appl	
27	43.8	1.2	5152	4	US-10-204-708-47	Sequence 47, Appl	

QY	1441	TTTCCATTGCCGTGACATCAGACACTCCAGCATCCAGGCATCTCTTGTCCTCCCAATAAC	1500
DB	1441	TTTCCATTGCCGTGACATGACAGACTCCAGCATCCAGGCATCTCTTGTCCTCCCAATAAC	1500
QY	1501	TGCTCTTAGATACATAGCCATATCTGTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAG	1560
DB	1501	TGCTCTTAGATACATAGCCATATCTGTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAG	1560
QY	1561	TTTTCTGGGAGGGTACACCCAAATCATCGATATCTTGTA TACTTTGAGCCCTTAGCGAC	1620
DB	1561	TTTTCTGGGAGGGTACACCCAAATGATCGATATCTTGTA TACTTTGAGCCCTTAGCGAC	1620
QY	1621	CTAACCAAATTTTAAAAATATCTTTTACCAGAAGTGTCTATTTCTCTGTA AAAACACTTTTT	1680
DB	1621	CTAACCAAATTTTAAAAATCTTTTACCAGAAGTGTCTATTTCTCTGTA AAAACACTTTTT	1680
QY	1681	TTTGGCAAGTTCACCTTTATCTTCCAATTAATTAATATATATTTATTTTAAATATTTT	1740
DB	1681	TTTGGCAAGTTCACCTTTATCTTCCAATTAATTAATATATTTATTTTAAATATTTT	1740
QY	1741	ATTTTCTTGCTAGGTATTAAGCTTTTGTAAATTAATTTTCAGTAGTCCCACCACTTCATA	1800
DB	1741	ATTTTCTTGCTAGGTATTAAGCTTTTGTAAATTAATTTTTCAGTAGTCCCACCACTTCATA	1800
QY	1801	GSTGGGAAGGATTTGGGGTTCTTCTGTGTGACGGGGCTGAAAATAACCAAGTGTCTCCAC	1860
DB	1801	GSTGGGAAGGATTTGGGGTTCTTCTGTGTGACGGGGCTGAAAATAACCAAGTGTCTCCAC	1860
QY	1861	CCTGCCACATATAGATGACAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGCTCCACTATC	1920
DB	1861	CCTGCCACATATAGATGACAGCCCATAGTTGGCCCCCTAGCTTCCAGCAGCTCCACTATC	1920
QY	1921	TGCCAGAGGAGCAAGGCTGCTTTAGACCTTAAGCCAGGGGGAAGAGCATCTTCATAAAAAA	1980
DB	1921	TGCCAGAGGAGCAAGGCTGCCWTAGACCTTAAGCCAGGGGGAAGAGCATCTTCATAAAAAA	1980
QY	1981	CTTTCAAGATCCAAAACAATTAATTTGTTTTTATTTATCTTGAGAAGTTCAGGCAATCAGT	2040
DB	1981	CTTTCAAGATCCAAAACAATTAATTTGTTTTTATTTATCTTGAGAAGTTCAGGCAATCAGT	2040
QY	2041	ATTTCCCAAGATGCGCACAAAGGCGACCAAGCGGCTTAGGATATCCAGCCTACCAAT	2100
DB	2041	ATTTCCCAAGATGCGCACAAAGGCGACCAAGCGGCTTAGGATATCCAGCCTACCAAT	2100
QY	2101	ATGCTCAATTCGACTTAATCAGAGGGTGAAGTGGCCCTGTCTCTTCTTTTCTTGGACCTC	2160
DB	2101	ATGCTCAATTCGACTTAATCAGAGGGTGAAGTGGCCCTGTCTCTTCTTTTCTTGGACCTC	2160
QY	2161	AGTTTCTCTCAGTGCTGGTAAGAAATGCATAACCTTTGATTTGATTAAGTTATAAATTC	2220
DB	2161	AGTTTCTCTCAGTGCTGGTAAGAAATGCATAACCTTTGATTTGATTAAGTTATAAATTC	2220
QY	2221	TGTGGTTCTGATCATTTGTCAGAGGGGAGATAGTTTCCGTGTCATTTTTCCTTCTCTCT	2280
DB	2221	TGTGGTTCTGATCATTTGTCAGAGGGGAGATAGTTTCCGTGTCATTTTTCCTTCTCTCT	2280
QY	2281	ATAGATAAATGAAATCTTGTTACTAGAACCAAGAAATGT CAGATGGCCAAAACAAGATG	2340
DB	2281	ATAGATAAATGAAATCTTGTTACTAGAACCAAGAAATGT CAGATGGCCAAAACAAGATG	2340
QY	2341	ACCAGATTTGATCTCAGCCTGATGACCCCTACAGCTCGTGTCTATGATATGAGTCCCTCATG	2400
DB	2341	ACCAGATTTGATCTCAGCCTGATGACCCCTACAGCTCGTGTCTATGATATGAGTCCCTCATG	2400
QY	2401	GGTAAAGCAGGGAAGAGTGGGAAAGAGAA CCA CCCA CTCTGCTTCATATTTGCAATTT	2460
DB	2401	GGTAAAGCAGGGAAGAGTGGGAAAGAGAA CCA CCCA CTCTGCTTCATATTTGCAATTT	2460
QY	2461	CATGTTTAACTCCCGCTGGAATATGAAGCATTCCTCTAGCATGAGGATPAAAGAAAG	2520
DB	2461	CATGTTTAACTCCCGCTGGAATATGAAGCATTCCTCTAGCATGAGGATPAAAGAAAG	2520





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

4 nucleic - nucleic search, using sw model

an on: March 11, 2004, 03:37:21 ; Search time 860.54 Seconds  
(without alignments)  
15799.545 Million cell updates/sec

itle: US-10-084-817-20

Effect score: 3693  
sequence: 1 CGATCGGCGGACTCCACCT.....ttaaagtgaaaaa 3693

oring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

sarched: 2432557 seqs, 1840798884 residues

otal number of hits satisfying chosen parameters: 4865114

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : Published Applications NA:\*\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09D\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	3693	100.0	3693	10	US-09-919-039-359
2	3693	100.0	3693	14	US-10-084-817-20
3	1940	52.5	1987	9	US-09-880-107-3287
4	1940	52.5	1987	10	US-09-960-706-941
5	1940	52.5	1987	10	US-09-873-319-617
6	1729.8	46.8	2655	13	US-10-098-841-144
7	1678.4	45.4	1755	9	US-09-764-877-3903
8	1678.4	45.4	1755	15	US-10-242-515-3903
9	630.4	17.1	634	9	US-09-738-973-540
10	630.4	17.1	634	9	US-09-854-133-540
11	630.4	17.1	634	14	US-10-144-649A-540
12	619.4	16.8	621	9	US-09-735-705-209
13	619.4	16.8	621	9	US-09-850-716A-209
14	619.4	16.8	621	9	US-09-897-778-209
15	619.4	16.8	621	10	US-09-466-396A-209

16	619.4	16.8	621	14	US-10-117-982-209	Sequence 209, App
17	619.4	16.8	621	15	US-10-313-986-209	Sequence 209, App
18	543.2	14.7	589	9	US-09-738-973-546	Sequence 546, App
19	543.2	14.7	589	9	US-09-854-133-546	Sequence 546, App
20	543.2	14.7	589	14	US-10-144-649A-546	Sequence 546, App
21	477.2	12.9	708	9	US-09-835-992A-34	Sequence 34, Appl
22	457.8	12.4	475	12	US-10-085-783A-56883	Sequence 56883, A
23	457.8	12.4	475	15	US-10-242-535A-56883	Sequence 56883, A
24	440.6	11.9	486	12	US-10-085-783A-56713	Sequence 56713, A
25	440.6	11.9	486	15	US-10-242-535A-56713	Sequence 56713, A
26	440.2	11.9	475	10	US-09-918-995-24257	Sequence 24257, A
27	437	11.8	462	12	US-10-085-783A-37434	Sequence 37434, A
28	437	11.8	462	15	US-10-242-535A-37434	Sequence 37434, A
29	434.6	11.8	504	10	US-09-918-995-25431	Sequence 25431, A
30	431.2	11.7	436	10	US-09-918-995-16944	Sequence 16944, A
31	427.6	11.6	435	10	US-09-918-995-8663	Sequence 8663, Ap
32	419.6	11.4	477	10	US-09-918-995-14933	Sequence 14933, A
33	401.4	10.9	415	12	US-10-085-783A-24844	Sequence 24844, A
34	401.4	10.9	415	15	US-10-242-535A-24844	Sequence 24844, A
35	398.2	10.8	416	12	US-10-085-783A-16890	Sequence 16890, A
36	398.2	10.8	416	15	US-10-242-535A-16890	Sequence 16890, A
37	385.8	10.4	390	12	US-10-085-783A-42307	Sequence 42307, A
38	385.8	10.4	390	15	US-10-242-535A-42307	Sequence 42307, A
39	381.8	10.3	410	12	US-10-085-783A-51574	Sequence 51574, A
40	381.8	10.3	410	15	US-10-242-535A-51574	Sequence 51574, A
41	375.2	10.2	395	12	US-10-085-783A-12488	Sequence 12488, A
42	375.2	10.2	395	15	US-10-242-535A-12488	Sequence 12488, A
43	372	10.1	372	12	US-10-085-783A-21666	Sequence 21666, A
44	372	10.1	372	15	US-10-242-535A-21666	Sequence 21666, A
45	356	9.6	356	14	US-10-116-712-45	Sequence 45, Appl

## ALIGNMENTS

### RESULT 1

US-09-919-039-359  
; Sequence 359, Application US/09919039  
; Publication No. US20030108871A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR FILING DATE: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 359  
; LENGTH: 3693  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1505038CB1  
US-09-919-039-359

Query Match	100.0%;	Score 3693;	DB 10;	Length 3693;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3693;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGATCGGCGGACTCCACCTCCGCTTACAGTCGTCGCGCGCTCTGCCCCGCGCCCC	60	
DB	1	CGATCGGCGGACTCCACCTCCGCTTACAGTCGTCGCGCGCTCTGCCCCGCGCCCC	60	
QY	61	AGGAGACCTGGACGACGACGATGGAAACGCTGGCTGGCGCTCGCGCTCGGTGT	120	
DB	61	AGGAGACCTGGACGACGACGATGGAAACGCTGGCTGGCGCTCGCGCTCGGTGT	120	
QY	121	GGCGGTGCTGGTCCGCGCCGCGGAGAGCTTAAGGCAATCCAAAGTCTGTGCCAA	180	
DB	121	GGCGGTGCTGGTCCGCGCCGCGGAGAGCTTAAGGCAATCCAAAGTCTGTGCCAA	180	



181 Y TGTGTTTGTGAGCGGCGCGGGAATGTGCACTCAAGAGAAAGGGAAACCCACCTGTCT 240  
181 b TGTGTTTGTGAGCGGCGCGGGAATGTGCACTCAAGAGAAAGGGAAACCCACCTGTCT 240  
241 Y CTGCAATTGAGCAATGCAAACTTCAAGAGAGGCTGTGTGTGCACTTAATGGCAAGACCTA 300  
241 b CTGCAATTGAGCAATGCAAACTTCAAGAGAGGCTGTGTGTGCACTTAATGGCAAGACCTA 300  
301 Y CTCAACCACTGTGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 360  
301 b CTCAACCACTGTGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 360  
361 Y TTACAGTGAACCTGCAAGAGAGGAAATCCAGTGAATGCAATGCAATGCAATGCAATGCA 420  
361 b TTACAGTGAACCTGCAAGAGAGGAAATCCAGTGAATGCAATGCAATGCAATGCAATGCA 420  
421 Y CTATCAGTCCAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480  
421 b CTATCAGTCCAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480  
481 Y CATCTCAGATGCTGCTTCTTAAAGGAGGAAATCCAGTGAATGCAATGCAATGCAATGCA 540  
481 b CATCTCAGATGCTGCTTCTTAAAGGAGGAAATCCAGTGAATGCAATGCAATGCAATGCA 540  
541 Y TAAAGAACTTGTGATGCTGCTTCTTAAAGGAGGAAATCCAGTGAATGCAATGCAATGCA 600  
541 b TAAAGAACTTGTGATGCTGCTTCTTAAAGGAGGAAATCCAGTGAATGCAATGCAATGCA 600  
601 Y ACAGAAATGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 660  
601 b ACAGAAATGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 660  
661 Y TAGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
661 b TAGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
721 Y CAGCTTCCAAAGATTTCTCAAGTGGCTCAAGCTGCTCAAGCTGCTCAAGCTGCTCAAG 780  
721 b CAGCTTCCAAAGATTTCTCAAGTGGCTCAAGCTGCTCAAGCTGCTCAAGCTGCTCAAG 780  
781 Y TGCCCTGGAGGATGAACCTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 840  
781 b TGCCCTGGAGGATGAACCTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 840  
841 Y TGTCTGTGCTGTGGAATTTGGGTCTGTACAGCTGATGCAATGCAATGCAATGCAATGCA 900  
841 b TGTCTGTGCTGTGGAATTTGGGTCTGTACAGCTGATGCAATGCAATGCAATGCAATGCA 900  
901 Y GGGGGCCAGACCCAGACAGAGAGGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 960  
901 b GGGGGCCAGACCCAGACAGAGAGGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 960  
961 Y TCAGAAACAGCTGAAAGCAAGAGAGTGAACCAAGAGATCTAATGAGGAGGAC 1020  
961 b TCAGAAACAGCTGAAAGCAAGAGAGTGAACCAAGAGATCTAATGAGGAGGAC 1020  
1021 Y AGACAGTCTGTGATCCAGAGATCTTCTCAAGTGAATGCAATGCAATGCAATGCAATGCA 1080  
1021 b AGACAGTCTGTGATCCAGAGATCTTCTCAAGTGAATGCAATGCAATGCAATGCAATGCA 1080  
1081 Y TGTCTGTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1140  
1081 b TGTCTGTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1140  
1141 Y TTAATTTGTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1200  
1141 b TTAATTTGTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1200  
1201 Y TTAATTTGTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1260  
1201 b TTAATTTGTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1260

1261 Y AAGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1320  
1261 b AAGAACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1320  
1321 Y ACTGCAATCTGTGAGAGGCTGCAAGAGGAGTGAACCAAGAGAGAGAGAGAGAGAGAG 1380  
1321 b ACTGCAATCTGTGAGAGGCTGCAAGAGGAGTGAACCAAGAGAGAGAGAGAGAGAGAG 1380  
1381 Y GTCCCGAGAGAGAGGTTGGGAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
1381 b GTCCCGAGAGAGAGGTTGGGAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
1441 Y TTTCCATGCTGCTGAGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
1441 b TTTCCATGCTGCTGAGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
1501 Y TGCTCTAGATACATAGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1501 b TGCTCTAGATACATAGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1561 Y TTTCTGGAGGAGTACACCCAAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
1561 b TTTCTGGAGGAGTACACCCAAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
1621 Y CTAAACCAAAATTTTAAATAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1621 b CTAAACCAAAATTTTAAATAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1681 Y TTTGGCAAGTGTGCTTATTTTAAATAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
1681 b TTTGGCAAGTGTGCTTATTTTAAATAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
1741 Y ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
1741 b ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
1801 Y GGTGGAAGAGAGTTGGGGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
1801 b GGTGGAAGAGAGTTGGGGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
1861 Y CTTGCCACATACCTAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
1861 b CTTGCCACATACCTAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
1921 Y TGCCAG 1980  
1921 b TGCCAG 1980  
1981 Y CTTTCAAGATCCAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2040  
1981 b CTTTCAAGATCCAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2040  
2041 Y ATTCCCAAG 2100  
2041 b ATTCCCAAG 2100  
2101 Y ATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
2101 b ATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
2161 Y AGTTTCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
2161 b AGTTTCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
2221 Y TGTGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
2221 b TGTGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
2281 Y ATAGAATAAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
2281 b ATAGAATAAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
2341 Y ACCAGATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400

2341 ACCAGATTGATCTCAGCTGATGACCTCAGCTCGTGTATGATATGAGTCTCATG 2400  
 2401 GGTAAGCAGGAGAGAGTGGGAAAGAGAACCCACCTCTGTCTTCATATTTGCAATTT 2460  
 2401 GGTAAGCAGGAGAGAGTGGGAAAGAGAACCCACCTCTGTCTTCATATTTGCAATTT 2460  
 2461 CATGTTAACTCCGCTGGGAAATAGAAAGCAATTCCTCTAGAGATGAGGATAAAGAAAG 2520  
 2461 CATGTTAACTCCGCTGGGAAATAGAAAGCAATTCCTCTAGAGATGAGGATAAAGAAAG 2520  
 2521 TTTCAGATTCAACAGGGGGAAGAAATGCGAGATTTAACTCCCTAAACTGTGACTTGGGGAG 2580  
 2521 TTTCAGATTCAACAGGGGGAAGAAATGCGAGATTTAACTCCCTAAACTGTGACTTGGGGAG 2580  
 2581 GTCACTCATTTACAGTTAGTCTGTGCTTTTTCGACTTCTGTGATTTAACTCCCTCACTC 2640  
 2581 GTCACTCATTTACAGTTAGTCTGTGCTTTTTCGACTTCTGTGATTTAACTCCCTCACTC 2640  
 2641 TACCTGTTTCAGATGCAATTTGGATATCAAGATTTAACTCCCTGACATAGATCTCATTT 2700  
 2641 TACCTGTTTCAGATGCAATTTGGATATCAAGATTTAACTCCCTGACATAGATCTCATTT 2700  
 2701 TGCAGAAAGCAGATTAAGACCATCAGAAAGAAATTTAGTGTGATGCAAGGCAA 2760  
 2701 TGCAGAAAGCAGATTAAGACCATCAGAAAGAAATTTAGTGTGATGCAAGGCAA 2760  
 2761 CTGTGAGAAACTGTTGTGCCAAATAGAAATTTAGTGTGATGCAAGGCAA 2820  
 2761 CTGTGAGAAACTGTTGTGCCAAATAGAAATTTAGTGTGATGCAAGGCAA 2820  
 2821 AAGGAGAAATTTCCACTTTGTTAGCAATTTCAAGCTTTTATGATPCCATCCATCTAAA 2880  
 2821 AAGGAGAAATTTCCACTTTGTTAGCAATTTCAAGCTTTTATGATPCCATCCATCTAAA 2880  
 2881 ACTCTTCAAACTCCACTGTTGCTGATGCAATTTCAAGCTTTTATGATPCCATCCATCTAAA 2940  
 2881 ACTCTTCAAACTCCACTGTTGCTGATGCAATTTCAAGCTTTTATGATPCCATCCATCTAAA 2940  
 2941 ACTCAGCAGCAGCAGCTTAACTAAAGTGTGTTTACAGGCTTTGACACTATGGGAGGAG 3000  
 2941 ACTCAGCAGCAGCAGCTTAACTAAAGTGTGTTTACAGGCTTTGACACTATGGGAGGAG 3000  
 3001 GCAAGAGTACA CAATTTGTTTAAAGCAAGAAACACAGTGTCTCTTCACTAGTCAATTA 3060  
 3001 GCAAGAGTACA CAATTTGTTTAAAGCAAGAAACACAGTGTCTCTTCACTAGTCAATTA 3060  
 3061 GCAATGTTTATCATCCAGACTTACTTACCTGCAACATTTCACTTCCCAAGAGCAATC 3120  
 3061 GCAATGTTTATCATCCAGACTTACTTACCTGCAACATTTCACTTCCCAAGAGCAATC 3120  
 3121 CACATTCCTCTTGTGAGTCTGCACTTCTGTGTAATAGGCGAGCTGTGCTATGCCGTA 3180  
 3121 CACATTCCTCTTGTGAGTCTGCACTTCTGTGTAATAGGCGAGCTGTGCTATGCCGTA 3180  
 3181 GAATCATGATCTGAGGACCATTCATGAGAGCTGTAATAGCTGCTGCTGCTGCTGCTGCT 3240  
 3181 GAATCATGATCTGAGGACCATTCATGAGAGCTGTAATAGCTGCTGCTGCTGCTGCTGCT 3240  
 3241 TCCATAAGTTTGTGATGAGCAAAACAAACAGGATTTAACTAGTGTGCTTCCCTCAGCC 3300  
 3241 TCCATAAGTTTGTGATGAGCAAAACAAACAGGATTTAACTAGTGTGCTTCCCTCAGCC 3300  
 3301 CTCTAAAGCATAGGCTTAGCTGAGGCTTCTTGGGCTTCTCTGTGTGTGTGTGTGTGTGT 3360  
 3301 CTCTAAAGCATAGGCTTAGCTGAGGCTTCTTGGGCTTCTCTGTGTGTGTGTGTGTGTGT 3360  
 3361 TGTAAACATATAGCATCTGTTAAGATCCAGTCTCCTGCAACCTTTCCCAATGCCGCTG 3420  
 3361 TGTAAACATATAGCATCTGTTAAGATCCAGTCTCCTGCAACCTTTCCCAATGCCGCTG 3420  
 3421 ACTCTGCACTATATCAGTTTTCGAAAGCAGGCTTCTCTGCTGTCTTAAACAGCCCACT 3480

3421 ACTCTGCACTATATCAGTTTTCGAAAGCAGGTTCTCTGCTGTCTTAAACAGCCCACT 3480  
 3481 GCAACAGTCTGAATGCTTCTCTTACCTTACCTATCTTCTTAACTAGTCTCAAGGAA 3540  
 3481 GCAACAGTCTGAATGCTTCTCTTACCTTACCTATCTTCTTAACTAGTCTCAAGGAA 3540  
 3541 CAATCTAAACAAAGTTTCTGTTGCAATGATGTTGTGAACTTGTATTTGTATTTAGTAGGC 3600  
 3541 CAATCTAAACAAAGTTTCTGTTGCAATGATGTTGTGAACTTGTATTTGTATTTAGTAGGC 3600  
 3601 TTCTATATTTGCAATTTAACTTCTTGTGAACTTGTATTTGTGAACTTGTATTTAGTAGGC 3660  
 3601 TTCTATATTTGCAATTTAACTTCTTGTGAACTTGTATTTGTGAACTTGTATTTAGTAGGC 3660  
 3661 ATGAATTAAGAAATTAAGTGAAGGAAAAA 3693  
 3661 ATGAATTAAGAAATTAAGTGAAGGAAAAA 3693

RESULT 2  
 US-10-084-817-20  
 ; Sequence 20, Application US/10084817  
 ; Publication No. US20030119009A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Susan Stuart  
 ; APPLICANT: Jed G. Nuchtern  
 ; APPLICANT: Sharon E. Plon  
 ; APPLICANT: Jason M. Shohet  
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
 ; FILE REFERENCE: PA-0046 US  
 ; CURRENT APPLICATION NUMBER: US/10/084,817  
 ; CURRENT FILING DATE: 2002-02-25  
 ; PRIOR APPLICATION NUMBER: 60/270,784  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 365  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 20  
 ; LENGTH: 3693  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 1505038CB1  
 US-10-084-817-20

Query Match 100.0%; Score 3693; DB 14; Length 3693;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGGCGGACTCCCACTCCGCTTACAGCTGCTGCGCGCTCTCTGCGCGCGCCCC 60  
 DB 1 CGATCGGCGGACTCCCACTCCGCTTACAGCTGCTGCGCGCTCTCTGCGCGCCCC 60  
 QY 61 AGGAGACTCGACACAGACACAGATGTGAAACGCTGCTGCGCTCGCGCTCGCGCTGCT 120  
 DB 61 AGGAGACTCGACACAGACACAGATGTGAAACGCTGCTGCGCTCGCGCTCGCGCTGCT 120  
 QY 121 GCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 121 GCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 181 TGTGTTTGTGAGCGCGCGCGGAAATGTGAGTCAAGAGAAAGGGGAAACCCACCTGTCT 240  
 DB 181 TGTGTTTGTGAGCGCGCGCGGAAATGTGAGTCAAGAGAAAGGGGAAACCCACCTGTCT 240  
 QY 241 CTGCAATTTAGCAATGCAAACTCACAAGAGGCTGTGTGTGCGAGTAATGGCAAGACTTA 300  
 DB 241 CTGCAATTTAGCAATGCAAACTCACAAGAGGCTGTGTGTGCGAGTAATGGCAAGACTTA 300  
 QY 301 CCTCAACCACTGTGAACTGCATCGAGATGCTGCTCCTCCTCCTGATCCAAATCCAGGTTGA 360  
 DB 301 CCTCAACCACTGTGAACTGCATCGAGATGCTGCTCCTCCTCCTGATCCAAATCCAGGTTGA 360

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1281	100.0	1281	14	US-10-084-817-1		Sequence 1, Appli
2	1255.8	98.0	1308	14	US-10-102-806-120		Sequence 120, App
3	398.6	31.1	475	10	US-09-918-995-23651		Sequence 23651, A
4	394.2	30.8	407	10	US-09-918-995-16444		Sequence 16444, A
5	105	8.2	469	14	US-10-198-846-11619		Sequence 11619, A
6	76.4	6.0	906	14	US-10-198-846-3058		Sequence 3058, Ap
7	60	4.7	60	10	US-09-908-975-8225		Sequence 8225, Ap
8	51	4.0	2100	12	US-10-482-1228-25713		Sequence 25713, A
9	50.6	4.0	2057	15	US-10-027-632-110305		Sequence 110305, A
10	50.6	4.0	2057	15	US-10-027-632-110306		Sequence 110306, A
11	48.2	3.8	500	10	US-09-991-936-3117		Sequence 3117, App
12	47.4	3.7	22693	12	US-10-471-416-6		Sequence 6, Appli
13	43.6	3.4	830	12	US-10-425-114-831		Sequence 831, App
14	42.2	3.3	671	14	US-10-184-644-346		Sequence 346, App
15	42.2	3.3	671	14	US-10-184-634-346		Sequence 346, App





2y 309 CCGCTGGAGCGCCCGCGCTGCTACTGCTGTTTCAGCGCGCAAGAGAAATCCGGGAAG 368  
Db 181 CCGCTGGAGCGCCCGCGCTGCTACTGCTGTTTCAGCGCGCAAGAGAAATCCGGGAAG 240  
2y 369 GACTTCGTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATCTGTGCTGCTCCCGG 428  
Db 241 GACTTCGTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATCTGTGCTGCTCCCGG 300  
2y 429 CTCTCTGCTCCACTCAAGGAACAGTATGCTTCAGAGCATGCTTGAACCTCCAGAGACTC 488  
Db 301 CTCTCTGCTCCACTCAAGGAACAGTATGCTTCAGAGCATGCTTGAACCTCCAGAGACTC 360  
2y 489 CTGGACACCCAGCAGCTTCAAGAGGCGCTTTTCGGAAGGACATGATCCG 535  
Db 361 CTGGACACCCAGCAGCTTCAAGAGGCGCTTTTCGGAAGGACATGATCCG 407

## RESULT 5

US-10-198-846-11619

; Sequence 11619, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11619

; LENGTH: 469

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1, 2, 3, 4

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-11619

Query Match 8.2%; Score 105; DB 14; Length 469;  
Best Local Similarity 95.6%; Pred. No. 4.6e-21;  
Matches 108; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 357 AAATCCGGGAAGGACTTCGTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATGCTGT 416  
Db 5 AAATCCGGGAAGGACTTCGTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATGCTGT 64

QY 417 GCTGTCCTCCGGCTCTCTGTCCTCACTCAAGGAACAGTATGCTCAGGAGCATGG 469  
Db 65 GCTGTCCTCCGGCTCTCTGTCCTCACTCAAGGAACAGTATGCTCAGGAGCATGG 117

## RESULT 6

US-10-198-846-3098/c

; Sequence 3098, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

US-10-198-846-3098/c

Query Match 4.7%; Score 60; DB 10; Length 60;  
Best Local Similarity 100.0%; Pred. No. 7.7e-08;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 GGATGTTCTCAGCCCTTGTCTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATACTAAT 1253  
Db 1 GGATGTTCTCAGCCCTTGTCTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATACTAAT 60

## RESULT 8

US-10-282-122A-25713

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3098

; LENGTH: 906

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1, 24, 346, 350, 495, 630, 632, 638, 640, 657, 663, 684,

; LOCATION: 691, 718, 721, 736, 755, 762, 766, 780, 817, 825, 826, 845,

; LOCATION: 849, 864, 879, 881, 894, 897, 904, 905

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-3098

## Query Match

6.0%; Score 76.4; DB 14; Length 906;

Best Local Similarity 76.8%; Pred. No. 1.9e-12;

Matches 119; Conservative 0; Mismatches 32; Indels 4; Gaps 2;

QY 319 GCGCCCCCGCGCTGCTACTGCTGTTTCAGCGCAAGAGAAATCCGGGA--GGACTTCGT 376  
Db 566 GCGGGCCGAGGCTACTTCTGTTTCAGCGCAAGAGAAATCCGGGAAGGACTTTCGT 507

QY 377 GACCA--GGCGCTGCAGAGCAGACTTGGAGCTGATGCTGCTGCTCCTCCGCTCTCT 434  
Db 506 GACCAAGCGCGNTGCAAGAGCAGACTTGGAGCTGATGCTGCTGCTCCTCCGCTCTCT 447

QY 435 GGTCCACTCAAGGAACAGTATGCTCAGGAGCATGG 469  
Db 446 GGTCCACTCAAGGAACAGTATGCTCAGGAGCATGG 412

## RESULT 7

US-09-908-975-8225

; Sequence 8225, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi

; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli

; APPLICANT: MINTZ, Liat

; APPLICANT: FAIGLER, Simchon

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE

; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME

; FILE REFERENCE: 3688-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8225

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-908-975-8225

Query Match 4.7%; Score 60; DB 10; Length 60;

Best Local Similarity 100.0%; Pred. No. 7.7e-08;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 GGATGTTCTCAGCCCTTGTCTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATACTAAT 1253  
Db 1 GGATGTTCTCAGCCCTTGTCTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATACTAAT 60



```

: Sequence 25713, Application US/10282122A
: Publication No. US20040029129A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangxu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: APPLICANT: Xu, H.
: TITLE OF INVENTION: Identification of E
: FILE REFERENCE: ELITRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,1
: CURRENT FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data remove
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 25713
: LENGTH: 2100
: TYPE: DNA
: ORGANISM: Mycobacterium avium
: IS-10-282-122A-25713

```

Query Match	4.0%	Score 51;	DB 12;	Length 2100;
Best Local Similarity	52.5%;	Pred. No. 9.4e-05;		
Matches 134;	Conservative 0;	Mismatches 120;	Indels 1;	Gaps 1;
2Y	121	GTTTGGACATGAGCTGAGGCGACGGTCTGGGGGGCGGTTCAGCCCTGTTCCGACGTACGGCGGA	180	
2b	1620	GTTCGCCAAGGAGCGGCGCGAAGAGGCCCTGTGAGGCGCGCAGGACGCGATGGCCCGCGA	1679	
2Y	181	GGAGGGGCGCGAATTGTTCTTGTGTCGGCTCCGCTTAGTGCCCGGTCCTCATTCGCGCGCGG	240	
2b	1680	G3TGCGCCGCGCGCGGGAATTCCGTTTGAGCGCTTGTGTCAAATGCGGAATCCA-TGAGCCGCGG	1738	
2Y	241	TGTCGCCGATTTTAGGGTAGGGAGAAGTGTCAGCTTTCAGGCATTCGCGAGCGGTGGCGGCC	300	
Db	1739	TGGCCGCGGAACCTGCATACGCCGACGTGTGCGCCCTCTACACCGCGATCGGTGAGGGGC	1798	
2Y	301	CCATGGCCCCCTGGGAGGCGCCCCCGCGGTGTGTACTGTGTTTCAGCGCGCAAGAGGAAAT	360	
Db	1799	ATGTGTCCGCCCGCGCATGTGTGCAGCGGCTGTGTGCCCGAGTTCGGCGCATTCGACCAGG	1858	
2Y	361	CCGGGAAGGACTTCG	375	
Db	1859	CCGAGGAGGACCTCG	1873	

```

RESULT 9
US-10-027-632-110305/c
/ Sequence 110305, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 110305
/ LENGTH: 2057
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(2057)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-110305

```

Query Match	4.0%;	Score 50.6;	DB 15;	Length 2057;
Best Local Similarity	40.0%;	Pred. No. 0.00012;		
Matches 206;	Conservative	0;	Mismatches 309;	Indels 0; Gaps 0;
QY	345	AGCGGCAAGAGGAAAATCCGGGAAGGACTTCGTGACCGAGGCGCTGCAGAGCAGCAGACTTGGGA	404	
Db	979	AGGGGCTGGAGAGGGGCTGGAGAGGAGGAGGCGCTGGAGAGGGCTGAAGGAGGAGGGG	920	
QY	405	GCTGATGCTGTGCTGCTCTCCGGCTCTCTGTGTCACACTCAAGGAAACAGTATGCTCAGGAG	464	
Db	919	CTCGATGCTGAGGCTCTGCGCAGCCGGAGAGGGGCGTGTGAGGAGGAGAGNNNNNNNN	860	
QY	465	CATGGCTTGAACTTCAGAGACTCCTGGACACAGCACCTACAAGGAGGGCCTTTCGGAAG	524	
Db	859	NN	800	
QY	525	GACATGATCGCTGGGGAGAGGAGAAACGCCAGGCTGACCCAGGCTTCTTTTGCAGGAAG	584	
Db	799	GGAGAGAGGGGCTTGGGGAGAGGAAAGGCGCTGGATGAGGAGAGGGCCCTGGAGGAGGAG	740	
QY	585	ATTGTGCGAGGGCATCTCCAGACCCATCTGGCTGCTGAGTGACACACAGGAGAGTGTCTGAC	644	
Db	739	GAGGGCTGGGAGAGAGAGACCTTGACGAGAAAGGAGGACCAGAGAGAGTGGAGGAG	680	
QY	645	ATCCAGTGGTTTCGGAGAGGCTATAGGGGCGGTGACGCAGACGGTCCCGCTTTAGCGTTG	704	
Db	679	AAGGAGGAGCAGCAGGAGGAGCTCGGGGGAAGAGAGGAGGAGCTAAGAAAGAAGGAGGAGATG	620	
QY	705	GAGCAGAGCCGACAGCAGCGGGGCTGGTGTTTCCAGCCAGGGGTGGACGATGCTGAGTCA	764	
Db	619	GAGGAGAGGGGCTTGATGGTGAAGGGCTGGGAAAGGAGAGGGGCTGGGAGCGGAGGTG	560	
QY	765	GAATGTGGGCTGACAACTTCGGGGACTTTTGACTGGGTCAATCGAGAAACCAATGGAGTTGAA	824	
Db	559	GGAGGAGGGGCTGGCAGAGGAGAAAGGGCTTGAGGAGGTACTGGAAGAAAGAAGAGGAGCAG	500	

2y 825 CAGCGCTGGAGGAGCAGTTGGAGAACCTGATAGA 859  
Db 499 GAGCTGGAGAGACGTTGGAGGAGGAGCTGGAGGA 465

## RESULT 10

US-10-027-632-110306/c  
; Sequence 110306, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; POLYMORPHISMS IN THE HUMAN GENOME  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 110306  
; LENGTH: 2057  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(2057)  
; OTHER INFORMATION: n = A, T, C or G

US-10-027-632-110306

Query Match 4.0%; Score 50.6; DB 15; Length 2057;  
Best Local Similarity 40.0%; Pred. No. 0.00012;  
Matches 206; Conservative 0; Mismatches 309; Indels 0; Gaps 0;  
Qy 345 AGCGCAAGAGGAATCCGGAAGGACTTCGTGACCGAGGCGCTGACAGCAGACTTGA 404  
Db 979 AGGGCTGGAGGAGGCGCTGGAAGGAGGAAGGGCTGGAGGAGGCGCTGAAGGAGGAGGG 920  
Qy 405 GCTGATGCTGTGCTGCTCCGCTCTCTGCTCCACTCAAGGAACAGTATGCTCAGGAG 464  
Db 919 CTCGATGCTGAGGCTCTGCGGAGCGGAGGAGGCGCTGTTGGAGGAGGAGNNNNNNNN 860  
Qy 465 CATGCTTGAACTTCCAGAGACTCTGTGACACAGCCTACAGGAGGCGCTTCCGAAG 524  
Db 859 NNN 800  
Qy 525 GACATGATCCGCTGGGAGAGGAGAAACCCAGGCTGACCCAGGCTTCTTTTCAGGAAG 584  
Db 799 CGAGGAGGAGGCGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 740  
Qy 585 ATTGTGGAGGAGCATCTCCAGCCCATCTGTGCTGTGTGAGTGACACAGGAGAGTCTTGAC 644  
Db 739 GAGGGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 680  
Qy 645 ATCCAGTGGTTCGGGAGGCGCTATGGGCGCTGACGAGCGGCTCCGCTGTGTAGCGTTG 704  
Db 679 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 620  
Qy 705 GAGCAGAGCCGACACAGCAGCGGGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 764  
Db 619 GAG 560

Qy 765 GAATGTGSCCTGGCAACTTCGGGACCTTTGACTGGGTCTATCGAGAACCATGGAGTTGAA 824  
Db 559 CGAGGAGGGGCTGGCAGAGGAGAGAGGGGCTGAGGAGTACTTGGHAAGAGAGAGGAGCAG 500  
Qy 825 CAGCGCTGGAGGAGCAGTTGGAGAACCTGATAGA 859  
Db 499 GAGCTGGAGGAGAGCAGTGGAGGAGGAGCTGGAGGA 465

## RESULT 11

US-09-991-936-317/c  
; Sequence 317, Application US/09991936  
; Publication No. US20030073827A1  
; GENERAL INFORMATION:  
; APPLICANT: Brandt, Kevin S.  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Stinchcomb, Dan T.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE  
; NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF  
; FILE REFERENCE: FC-6-C1  
; CURRENT APPLICATION NUMBER: US/09/991.936  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US/09/543.668  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: 60/128,704  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 1959  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 317  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Ctenocephalides felis  
; US-09-991-936-317

Query Match 3.8%; Score 48.2; DB 10; Length 500;  
Best Local Similarity 48.4%; Pred. No. 0.00045;  
Matches 119; Conservative 0; Mismatches 127; Indels 0; Gaps 0;  
Qy 566 AGGCTTCTTTTCAGGAGATGTGGAGGCGATCTCCAGGCCATCTCGCTGGTGAGTGA 625  
Db 441 ATGGTTATTTTGTAAAGGGCTTCCAAAGCAATTTCAAACCTATTTCATAGTTAGTGA 382  
Qy 626 CACACGGAGAGTGTCTGACATCCAGTGTGTTTCGGAGGCGCTATGGGCGGTGACGAGAC 685  
Db 381 TATCCGAGGAACNAGATATAAAATGTTTATTGAAAAATATGCTGCTAAAGTTATAAC 322  
Qy 686 GGTCCGCTGTGAGGCTTGGAGCAGAGCCGAGCAGGCGGCTGGTGTTCAGCCAGG 745  
Db 321 NATTGAATTAATTCAGTAATGCAATTCGANAATNTCGAGTTGAAAATTTACNAAAG 262  
Qy 746 GGTGACGATGCTGAGTCAGATGTGGCTGGCAACTTCGGGAGCTTTGACTGGGTCTAT 805  
Db 261 TCTTGATGATGTTGCTNTGATGATTATATATATATATATATATATATATATATAT 202  
Qy 806 CGAGAA 811  
Db 201 TGTAA 196

## RESULT 12

US-10-271-416-6  
; Sequence 6, Application US/10271416  
; Publication No. US20040043021A1  
; GENERAL INFORMATION:  
; APPLICANT: Keith, Tim  
; APPLICANT: Little, Randall D.  
; APPLICANT: Van Berdewegh, Paul  
; APPLICANT: Dupuis, Josee  
; APPLICANT: Del Masaro, Richard G.  
; APPLICANT: Allen, Kristina  
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES



717	QY	CAGCAGCGGGCTGGGTGTTACGCCACGAGGGGTGACAGATGCTGAGTCAGAAATGTGGCCTG	716
306	Db	DSDEVDIRSEKWEDEARRRELEARRRQEELERLREQEKEKERRERADRGEAERG	365
777	QY	GACAACTCGGGGACTTTGACTGGGTCTATCGAGAACCATGGAGTTGAACAGCGCCTGGAG	836
366	Db	SGGSSGDELREDDPEVKRGRKRGRRGPPSSSDSEPEALEREAKSKAKPPSSSTFPAR	425
837	QY	GAGCAGTTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTTACTACTAGTTCCTAGG	896
426	Db	KPG--QKEKVRPEEKQAKPKVKVETRRKRSSEGFMSMDRKVEKKKEPSEVEKLQLHSEIK	483
897	QY	AGTACGCTGGGGCTGCTGAGTGGGGGTGGGGCTGACTCTGCAAAATGGGGGTGTCGCC	956
484	Db	FALKVDSFDVKRCLNAALEELGTLQVTSQILQKNTDVATLKKIRRYKANKOVMEKAAEVI	543
957	QY	CGATCCTGGCGAGGTGAGGAACAGACAGCGGGGGGTCTAGATTCTTGAGGGGTGTGGTGA	1016
544	Db	TRLKSRVLPFKIEAVQVKNKAGMEKEAKEELAGELAGEEAPQEKADKEDSTDLSPVN	603
1017	QY	TATTGGG	1023
604	Db	GEATSOK	610

Search completed: March 12, 2004, 06:13:07  
Job time : 307.609 secs

Search completed: March 12, 2004, 06:13:07  
Job time : 307.609 secs

result No.	Score	Query Match	length	DB	ID	Description	
1	317.2	24.8	419	4	US-09-621-976-3203	Sequence 3203, Appl	
C	2	3.49	3.8	7218	1	US-08-2332-463-14	Sequence 14, Appl
C	3	39.8	3.1	4403765	3	US-09-103-8403-2	Sequence 2, Appl
C	4	39.8	3.1	4411529	3	US-09-103-8403-1	Sequence 1, Appl
C	5	39.6	3.1	1336	4	US-09-252-91A-7459	Sequence 7459, Appl
C	6	39.6	3.1	1665	4	US-09-252-91A-7242	Sequence 7242, Appl
C	7	39.6	3.1	1713	4	US-09-252-91A-7383	Sequence 7383, Appl
C	8	39.2	3.1	834	4	US-09-621-976-2574	Sequence 2574, Appl
C	9	38.8	3.0	1644	4	US-09-252-91A-6420	Sequence 6420, Appl
10	38.8	3.0	2088	4	US-09-252-91A-6339	Sequence 6339, Appl	
C	11	38.8	3.0	2202	4	US-09-252-91A-6131	Sequence 6131, Appl
C	12	37.2	2.9	1437	4	US-09-647-224A-9	Sequence 9, Appl
C	13	37	2.9	1036	4	US-09-072-596-304	Sequence 304, Appl
C	14	37	2.9	1036	4	US-09-072-967-309	Sequence 309, Appl
C	15	36.4	2.8	831	4	US-09-252-91A-532	Sequence 532, Appl
C	16	36.4	2.8	2478	4	US-09-252-91A-592	Sequence 592, Appl
C	17	36	2.8	4078	4	US-09-016-434-1120	Sequence 1120, Appl
C	18	35.6	2.8	314	4	US-09-615-510A-8	Sequence 8, Appl
C	19	35.6	2.8	430	4	US-09-621-976-16656	Sequence 16656, Appl
C	20	35.6	2.8	2291	2	US-08-476-062A-53	Sequence 53, Appl
C	21	35.6	2.8	2310	2	US-08-476-062A-41	Sequence 41, Appl
C	22	35.6	2.8	2310	5	PCT-US96-01314-41	Sequence 41, Appl
C	23	35.6	2.8	2405	3	US-08-549-846-3	Sequence 3, Appl
C	24	35.6	2.8	2776	4	US-09-023-655-1146	Sequence 1146, Appl
C	25	35.2	2.7	1305	4	US-09-252-91A-5951	Sequence 5951, Appl
C	26	35	2.7	2511	4	US-09-489-039A-530	Sequence 530, Appl
C	27	34.8	2.7	1073	4	US-09-072-596-322	Sequence 322, Appl

589 TGGAGGCACTCCAGCCCATCTGGCTGGTGTGACACACGAGAGTGTCTGACATCC 648  
295 TGGAGGCACTCCAGCCCATCTGGCTGGTGTGACACACGAGAGTGTCTGACATCC 354  
649 AGTGGTTGGGAGGCTATGGGCGCTGACGACAGGTCGGGTTGTAGCGTTGGAGC 708  
355 AGTGGTTGGGAGGCTATGGGCGCTGACGACAGGTCGGGTTGTAGCGTTGGAGC 413  
709 AGAGCC 714  
414 AGAGCC 419

RESULT 2  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/232.463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935.313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 INMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ9pt-Fls  
US-08-232-463-14  
Query Match 3.8%; Score 49; DB 1; Length 7218;  
Best Local Similarity 3.3%; Pred.No. 0.002;  
Matches 13; Conservative 218; Mismatches 158; Indels 0; Gaps 0;

651 TGGTTCCGGAGGCGCTATGGGCGCTGACGACAGCGTCCGGTTGTAGCGTTGAGCAG 710  
1440 TGTACRR 1381  
711 AGCCGACACAGCGGGGCTGGTGTTCACGCCAGGGGTGGACGATGCTGAGTCAGATGT 770

Db 1380 RRR 1321  
QY 771 GCGCTGGACAACCTTGGGCACTTTGACTGGGTATCGAGAACCATGGAGTTGAACAGCGC 830  
Db 1320 RRR 1261  
QY 831 CTGGAGGAGCAGTTGGGAGAACCTGATAGATTTATCCGCTCCAGACTTTACTCACTAGGT 890  
Db 1260 RRR 1201  
QY 891 TCTAGGAGTGGGCTGCTGAGTGGGCTGGGCTGACTCTGCAAAATGGGGGT 950  
Db 1200 RRR 1141  
QY 951 GTCCCCCGATCTGGCCGAGGTGAGGACACACAGGGGGGTCTAGATTCTGAGGGGGTT 1010  
Db 1140 RRR 1081  
QY 1011 GGTGGATATTGGGCAAGCAGCAAACTTC 1039  
Db 1080 RRR 1052

RESULT 3  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2  
Query Match 3.1%; Score 39.8; DB 3; Length 4403765;  
Best Local Similarity 49.8%; Pred.No. 5.9;  
Matches 127; Conservative 0; Mismatches 127; Indels 1; Gaps 1;

QY 121 GTTGGACATGAGCTGAGGGCAGCGTCCGGCGCTCAGCCCTGTCGACGCTACGGCA 180  
Db 2904687 GTTCCCAAGGAGCGCGCTGAGAGCGCTTGGAGACCGGTAAGGATGGATGGCCCGCA 2904628  
QY 181 GGAGGGCGCGATTGTTCTTGTTCGCTCCGCTTAGTGGCGCGCTCCATTCCCGCGG 240  
Db 2904627 GTTCCCGCGCGTGGACTTCCGTTGCGCGCTTGGTCAATGGTGTTCATGGCG-GCGG 2904569  
QY 241 TGTCCCGATTGTAGGGGTAGGAGAGTGTCTAGCTTACGACATCCGAGGCGTGGCGCC 300  
Db 2904568 TGGCCCGCGAGCTGCACTACGCGGACGCTGTTCAGCACTCTATACCGCCATTCGGTGGAGGCG 2904509  
QY 301 CCATGGCCCGCTGGGAGCGCCCGCGCTGTCTGTCTGCTGACGCGCAAGAGAAAT 360  
Db 2904508 ACCTGTCCGGGAAACACGCTGCTGACGCGGTTGTGGCCGAGTCTCGCGGTATCGACCAGG 2904449  
QY 361 CCGGGAAGGACTTCG 375  
Db 2904448 CGGAAGGGAACTCG 2904434



Matches 105; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 307 CCCCGCTGGAGGCGCCCGCGGTGGTACTGCTTTTCAGCGGCAGAGAAATCCGGGA 366  
Db 830 CCGCGCGCGGTTCACAGATGCCATTTCGTCTGTCGGCGCCTCACCGGTTGCGCA 889  
QY 367 AGGACTTCGTGACCAGGCGCTGCAGAGCAGACTTGGAGCTGATCTGTGTGCTCTCTCC 426  
Db 890 AGACCGAGTGATCGCGGCTGGACACACCTCGATCTCGAAGCCACGCCAACCATC 949  
QY 427 GCCTCTCTGTGTCCTCACTCAAGAACAGTAGTCTCAGGAGCATGGCTTGAACCTCCAGAC 486  
Db 950 GCGGCTCCAGCTTTGGCGCGGAGCCACGCCGACCGCGCAGATCGACTTCGAGAACC 1009  
QY 487 TCTCGACACAGCACCTTACAAGGAGGCGCTTTTCG 520  
Db 1010 GCTGGCCATCGACATCTCGAAGAGCGTCATCG 1043

RESULT 6  
US-09-252-991A-7242/c  
; Sequence 7242, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7242  
; LENGTH: 1665  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7242

Query Match 3.1%; Score 39.6; DB 4; Length 1665;  
Best Local Similarity 49.1%; Pred: No.0.41;  
Matches 105; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 307 CCCCGCTGGAGGCGCCCGCGGTGGTACTGCTTTTCAGCGGCAGAGAAATCCGGGA 366  
Db 1064 CCGCGCGCGGTTCACAGATGCCATTTCGTCTGTCGGCGCCTCACCGGTTGCGCA 1005  
QY 367 AGGACTTCGTGACCAGGCGCTGCAGAGCAGACTTGGAGCTGATCTGTGTGCTCTCTCC 426  
Db 1004 AGACCGAGTGATCGCGGCTGGACACACCTCGATCTCGAAGCCACGCCAACCATC 945  
QY 427 GCCTCTCTGTGTCCTCACTCAAGAACAGTAGTCTCAGGAGCATGGCTTGAACCTCCAGAC 486  
Db 944 GCGGCTCCAGCTTTGGCGCGGAGCCACGCCGACCGCGCAGATCGACTTCGAGAACC 885  
QY 487 TCTCGACACAGCACCTTACAAGGAGGCGCTTTTCG 520  
Db 884 GCTGGCCATCGACATCTCGAAGAGCGTCATCG 851

RESULT 7  
US-09-252-991A-7383  
; Sequence 7383, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18

[illegible]

39 GGGTCGGCAACGTCCTCCGGTGGCGCGGACGCCGATCGCGGTGCAGAGCATGACCAACA 98  
355 GGAATTCGGGGAAGGACTCTGTGACCGAGGCGCTGCAGAGCAGACTTGGAGCTGATGTCT 414  
99 CCGAGACCTCGAAGTCGCTCGCTGCCACCGTTCGGCGCAGATCCGGCGCTTGGAGATCGCGGC 158  
415 GTGCTGTCTCCGGCTCTCTGTGCTCAACGA 448  
159 CGGACATCGTGGGGTTTCGGTCCCGGACATGGA 192

```

1      RESULT 11
2      JS-09-252-991A-6131/c
3      ; Sequence 6131, Application US/09252991A
4      ; Patent No. 6551795
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Marc J. Rubenfield et al.
7      ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
8      ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
9      ; FILE REFERENCE: 107196.136
10     ; CURRENT APPLICATION NUMBER: US/09/252,991A
11     ; CURRENT FILING DATE: 1999-02-18
12     ; PRIOR APPLICATION NUMBER: US 60/074,788
13     ; PRIOR FILING DATE: 1998-02-18
14     ; PRIOR APPLICATION NUMBER: US 60/094,190
15     ; PRIOR FILING DATE: 1998-07-27
16     ; NUMBER OF SEQ ID NOS: 33142
17     ; SEQ ID NO 6131
18     ; LENGTH: 2202
19     ; TYPE: DNA
20     ; ORGANISM: Pseudomonas aeruginosa
21     JS-09-252-991A-6131

```

Query Match	3.0%;	Score 38.8;	DB 4;	Length 2202;
Best Local Similarity	53.2%;	Pred. No. 0.75;		
Matches	82;	Conservative 0;	Mismatches 72;	Indels 0; Gaps 0;
2y	295	GCGCCCATGGCCCGCTGGAGGCGCCCGCGCTGGTACTGCTTTCAGCGGCAAGA	354	
3b	2081	GGGTGGCAACGTCCCGGTGGCGGCGACGCGCGCATCGCGTGCAGAGCATGACCAACA	2022	
2y	355	GGAAATCCGGGAAGGACTTGTGTACCGAGGCGCTGCAGAGCAGACTTGGAGCTGATGTCT	414	
3b	2021	CCGNAGACTCGGACGTGCTTGCCACCGTTCGGCAGATCCGCCGCTCGAAGATGCCGGCG	1962	
2y	415	GTGCTGTCTCCCGCTCTCTGGTCCACTCAAGGA	448	
3b	1961	CGGACATCGTGGGGTTTCGTCCCGCACATGGA	1928	

RESULT 12  
US-09-647-224A-9  
Sequence 9, Application US/09647224A  
Patent No. 6482631  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Guttridge, Steven  
APPLICANT: Hitz, William D.  
APPLICANT: Maxwell, Carl A.  
APPLICANT: Rafalski, J. Antoni  
APPLICANT: Tso, Yong  
APPLICANT: Vollmer, Steven J.  
TITLE OF INVENTION: TRYPTOPHAN BIOSYNTHETIC ENZYMES  
CURRENT APPLICATION NUMBER: US/09/647,224A  
CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/079,386  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: PCT/US99/05046  
PRIOR FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 32

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1437
; TYPE: DNA
; ORGANISM: Zea mays
US-09-647-224A-9

Query Match          2.9%; Score 37.2; DB 4; Length 1437;
Best Local Similarity 44.7%; Pred. No. 1.7;
Matches 144; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Qy      746  GTGTGGAGCATGCTGAGTCAGAAATGTGGCCTGGAGCAACTTCGGGGACATTCGACTGGGTCAT 805
Db      363  GGAGATAGAGGAGGAGCTGGGAGCCCCCTTAACATCCTAGAGAAGATCATTTTGGGACAA 422

Qy      806  CGAGAACCATGTGAGTTCGAACAGCGCCCTGGAGGAGCAGTTTGAGAGAACCTGATAGAAATTTAT 865
Db      423  GGAGATTGAAGTAGCTGAGGGGCATGCTAAGAAGCCCTTTGAGGAGGTGATTCAGGCTGC 482

Qy      866  CCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCCTCTCAGGTTGGGGGT 925
Db      483  AAGCAAAGCCCCCTCTTCAAGAGACTTCTATTGGCGCTTTTAAAGCTGCCCTACAAGCGTAA 542

Qy      926  GGGGCTGACTCTGC AAAATGGGGGTGTCCCCCGATCTGCGCCGAGGTGAGGAAACAGACAG 985
Db      543  TGGGGTGCCTGCATTGATTGCTGAGGTCAAGAAAGCATCCCCGAGTAGGGGTGTGCTCAG 602

Qy      986  GGGGGGTCTAGATTCCTGAGGGGTGTGGTGATATTGGGCAAGGCAGAGAAACCTCTGGAGA 1045
Db      603  GGAAGACTTTAAATCCCTGTGAATTTGCTCAAGCTTATGAAGAAGATCGAGCTGCATGTTT 662

Qy      1046  CCTCATTTTCTCCATGGGGAAG 1067
Db      663  GAGCATTTTCACAGATGAGGAAG 684

```

RESULT 13  
US-09-072-596-304/c  
; Sequence 304, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; TUBERCULOSIS  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 304:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1036 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-072-596-304

Query Match 2.9%; Score 37; DB 4; Length 1036;  
Best Local Similarity 29.3%; Pred. No. 1.7;  
Matches 118; Conservative 55; Mismatches 230; Indels 0; Gaps 0;  
QY 613 GCCTGGTGAAGTACACACGAGAGTGTCTGACATCCAGTGTTCGGAGGCTATGGGG 672  
DB 998 GGAGANYGNGGAAGAGAGAMWGMKNTNRNGAKAANRNGWGMWAMCSGRAGAMKVXGKGG 939  
QY 673 CCGTGACGACAGCGTCCGCGTGTAGCGTTGGAGCAGAGCCGACAGCGGGGCTGGG 732  
DB 938 GAGGGGGGAAAGGKWMNRVMMWGGGGGRTNKAKAGGGGSGMGRGWAMSGMGGKRVGGG 879  
QY 733 TGTTCACCCAGGGGTGACAGTCTGAGTCAGAAATGTGGCCTGGACAACTTCGGGGACT 792  
DB 878 WKTAAANWNGNGGGNRNGAGWGTCTATWTKWNGWMMNRWTRWGGGTTTTATAMKT 819  
QY 793 TTGACTGGGTATCATGAGAACCATGAGTTGAACAGCGCTGGAGGACAGTTGGAGAAC 852  
DB 818 GASASMGGGGAGYGGGGYGNARTTTGGMAGKANAAMNSNGMKGWANNATTTGGNAKS 759  
QY 853 TGATAGATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTG 912  
DB 758 CMRGMGGGTNGTIMGAGNGNNGWKTSMGAGGGGNGMKYARGGGGAGNAAAGCCY 699  
QY 913 CTGAGTGGGGTGGGGTGTACTCTGAAAATGGGGTGTCCCCCGATCTGGCCGAGGT 972  
DB 698 GNMWTTARSGGGGNGNSANAWGGGGGTTAAACCGGGGAGKTTNAAGGGGKGGGGAAACMC 639  
QY 973 GAGGAACAGACAGGGGGGCTCTAGATTCTGAGGGGGTGTGGTG 1015  
DB 638 GGRASATAANNMGGGGGGGGSNANTTCGNGGAGTNSAAGS 596

RESULT 14  
US-09-072-967-309/c  
Sequence 309, Application US/09072967  
Patent No. 6592877  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 304:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1036 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-072-596-304

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072.967  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 309:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1036 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-072-967-309

Query Match 2.9%; Score 37; DB 4; Length 1036;  
Best Local Similarity 29.3%; Pred. No. 1.7;  
Matches 118; Conservative 55; Mismatches 230; Indels 0; Gaps 0;  
QY 613 GCCTGGTGAAGTACACACGAGAGTGTCTGACATCCAGTGTTCGGAGGCTATGGGG 672  
DB 998 GGAGANYGNGGAAGAGAGAMWGMKNTNRNGAKAANRNGWGMWAMCSGRAGAMKVXGKGG 939  
QY 673 CCGTGACGACAGCGTCCGCGTGTAGCGTTGGAGCAGAGCCGACAGCGGGGCTGGG 732  
DB 938 GAGGGGGGAAAGGKWMNRVMMWGGGGGRTNKAKAGGGGSGMGRGWAMSGMGGKRVGGG 879  
QY 733 TGTTCACCCAGGGGTGACAGTCTGAGTCAGAAATGTGGCCTGGACAACTTCGGGGACT 792  
DB 878 WKTAAANWNGNGGGNRNGAGWGTCTATWTKWNGWMMNRWTRWGGGTTTTATAMKT 819  
QY 793 TTGACTGGGTATCATGAGAACCATGAGTTGAACAGCGCTGGAGGACAGTTGGAGAAC 852  
DB 818 GASASMGGGGAGYGGGGYGNARTTTGGMAGKANAAMNSNGMKGWANNATTTGGNAKS 759  
QY 853 TGATAGATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTG 912  
DB 758 CMRGMGGGTNGTIMGAGNGNNGWKTSMGAGGGGNGMKYARGGGGAGNAAAGCCY 699  
QY 913 CTGAGTGGGGTGGGGTGTACTCTGAAAATGGGGTGTCCCCCGATCTGGCCGAGGT 972  
DB 698 GNMWTTARSGGGGNGNSANAWGGGGGTTAAACCGGGGAGKTTNAAGGGGKGGGGAAACMC 639  
QY 973 GAGGAACAGACAGGGGGGCTCTAGATTCTGAGGGGGTGTGGTG 1015  
DB 638 GGRASATAANNMGGGGGGGGSNANTTCGNGGAGTNSAAGS 596

RESULT 15  
US-09-252-991A-532/c  
Sequence 532, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 532  
LENGTH: 831  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa

JS-09-252-991A-532

Query Match 2.8%; Score 36.4; DB 4; Length 831;  
 Best Local Similarity 53.5%; Pred. No. 2.3;  
 Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

2y 689 CCGCGTTGTAGCGTTGAGCAGAGCCGACAGCAGCGGGGCTGGGTGTTCACGCCAGGGGT 748  
 |||||  
 289 CCACCTTCGTTTCGCCGATACCGAGCCGAGCGGCTGGGGCTTCGCGATCTTCGACCGGCA 230  
 |||||

2y 749 GGACGATGCTGAGTCAGAAATGTGGCCTGGACAACTTCGGGACTTTGACTGGGTGATCGA 808  
 |||||

2b 229 GGAGCGGTTCGTTTACCGCTGTGCCGTGGACAACTTCTGGGAGCTGGACAGCGCTACCGA 170  
 |||||

2y 809 GAACCATGGAGTTGAACAGGCC 830  
 |||||

2b 169 GCGGAGGTGATCGCGCGGCAC 148  
 |||||

Search completed: March 11, 2004, 14:41:22  
 Job time : 76.4979 secs





1	1255.8	98.0	1308	3	AAF21733	Human bre
2	459.2	35.8	465	3	AC02135	Human sec
3	398.6	31.1	475	8	ACh36439	Human end
4	394.2	30.8	407	8	ACh29232	Human adu
5	292.6	22.8	337	2	AAT23873	Human gen
6	125	9.8	502	4	AA120639	Human bre
7	82	6.4	570	4	ABL20639	Human bre
8	76.4	6.0	568	4	ABL30119	Drosophil
9	60.8	4.7	2673	4	ABL11749	Human bre
10	60.8	4.7	3406	4	ABL30184	Drosophil
11	60	4.7	60	6	ABN35477	Human spl
12	59.4	4.6	2000	7	ADA71938	Rice gene
13	56.8	4.4	467	9	ADB51434	Primary r
14	51	4.0	2100	7	ACA37843	Primary r
15	48.8	3.8	500	3	ACA93822	Cat flea
16	43.2	3.4	3704	4	ABL30114	Drosophil
17	38.2	3.1	2373	7	ACA40692	Prokaryot
18	39.8	3.1	2379	7	ACA38426	Prokaryot
19	39.8	3.1	80557	6	ABX09142	Mycobacte
20	39.8	3.1	110000	4	AA199682	Continuation (29 0
21	39.8	3.1	110000	4	AA199682	Continuation (30 0
22	39.8	3.1	110000	4	AA199683	Continuation (29 0
23	39.8	3.1	110000	4	AA199683	Continuation (30 0

sequences AAF22032 - AAF22040 and AAB59129 which are used in the isolation and characterisation of the DNA and protein sequences of the invention. The breast and ovarian cancer associated DNA, protein, agonist or antagonist sequences exhibit cytostatic, immunosuppressive, neurotropic, neuroprotective, antiviral, antiallergic, hepatotropic, antidiabetic, antinflammatory, antitumor, anticonvulsant, antibacterial, antifungal, antiparasitic and cardiac activity. The polynucleotide and protein sequences are used in the diagnosis of cancer, particularly breast and ovarian cancer. The nucleic acid sequences, proteins, agonists and antagonists may also be used in the diagnosis, prevention and treatment of immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular diseases such as myocardial ischaemia; wound healing; neurological diseases such as cerebral anoxia and epilepsy; and infectious diseases

Sequence 1308 BP; 280 A; 325 C; 433 G; 269 T; 0 U; 1 Other;

Query Match 98.0%; Score 1255.8; DB 3; Length 1308;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1268; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

9 CGGGCTCGACTGTCTTAAGTGAAGTTCGGGTGGGGAGCTTCAGAGGGAGGCTGCTCT 68  
9 CGGTCGCGACTGTCTTAAGTGAAGTTCGGGTGGGGAGCTTCAGAGGGAGGCTGCTCT 68  
69 GTGAAGGAACCGCTTCTCTCCGGGTGTCTACCGCTTTCTCCCATATCTGTTGGAC 128  
69 GTGAAGGAACCGCTTCTCTCCGGGTGTCTACCGCTTTCTCCCATATCTGTTGGAC 128  
129 ATGAGCTCAGAGGACGGTTCGGGGCGGTTCAGCTTTCGGAGCTACGGGAGAGGGGC 188  
129 ATGAGCTCAGAGGACGGTTCGGGGCGGTTCAGCTTTCGGAGCTACGGGAGAGGGGC 187  
189 CGGATGTTCTTGTTCGGCTTCGGCTTGTAGTGGCGGTTCATTCGGCGGGTGTCCCGA 248  
188 GCGATTGTTCTTGTTCGGCTTCGGCTTGTAGTGGCGGTTCATTCGGCGGGTGTCCCGA 247  
249 TTTTAGGGGTAGGAGAGTGTACGCTTCAGGATCGGGGGCGTGGCGGCCCATGGCC 308  
248 TTTTAGGGGTAGGAGAGTGTACGCTTCAGGATCGGGGGCGTGGCGGCCCATGGCC 307  
309 CGGCTGGAGGGCGCCCGCGGTGTACTGTGTTTCAGCGCAAGAGAAATCCGGGAAG 368  
308 CGGCTGGAGGGCGCCCGCGGTGTACTGTGTTTCAGCGCAAGAGAAATCCGGGAAG 367  
369 GACTTCGTGACCGAGCGCTTCAGAGCAGACTTCGGAGCTGATGCTGTGCTGCTCCGG 428  
368 GACTTCGTGACCGAGCGCTTCAGAGCAGACTTCGGAGCTGATGCTGTGCTGCTCCGG 427  
429 CTCTCTGTTCCACTCAAGGAACAGTATGCTCAGGAGCATGCTTGAACCTCCAGAGACTC 488  
428 CTCTCTGTTCCACTCAAGGAACAGTATGCTCAGGAGCATGCTTGAACCTCCAGAGACTC 487  
489 CTGGACACAGACCTACAGAGAGCGCTTCGGAAGCAGATGATCCGCTGGGAGAGGAG 548  
488 CTGGACACAGACCTACAGAGAGCGCTTCGGAAGCAGATGATCCGCTGGGAGAGGAG 547  
549 AAACGCCAGGCTGACCCAGGCTTTCTTTTCAGGAAGATTGCGAGGGCATCTCCAGGCC 608  
548 AAACGCCAGGCTGACCCAGGCTTTCTTTTCAGGAAGATTGCGAGGGCATCTCCAGGCC 607  
609 ATCTGGCTGGTGAAGTGAACACGGAGAGTGTCTGATCCAGTGGTTTCGGAGGCTAT 668  
608 ATCTGGCTGGTGAAGTGAACACGGAGAGTGTCTGATCCAGTGGTTTCGGAGGCTAT 667  
669 GGGGGCGTCAACGACAGCGGTTCGGGTGTAGGCTTTGGAGCAGAGCCGACAGCGGGGC 728  
668 GGGGGCGTCAACGACAGCGGTTCGGGTGTAGGCTTTGGAGCAGAGCCGACAGCGGGGC 727  
729 TGGGTGTTCAACGACAGCGGTTCGGGTGTAGTGTGATGTCAGAAATGTGGCTCGAACACTCGG 788

728 TGGGTGTTCAACGACAGCGGTTCGGAGTGTGAGTCAAGATGCGCTGGACAACACTCGG 787  
789 GACTTTGACTGGGTTCATCGAAGCAATGGAGTTGAACAGAGCGCTGGAGGAGCAGTTGGAG 848  
798 GACTTTGACTGGGTTCATCGAAGCAATGGAGTTGAACAGAGCGCTGGAGGAGCAGTTGGAG 847  
849 AACCTGTAGTAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAAGCTGGG 908  
848 AACCTGTAGTAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAAGCTGGG 907  
909 CCGTGTGAGGTGGGGGTGGGGGTGACTCTGCAAAATGGGGGTGTCCTCCGATCTCTGGCG 968  
908 CCGTGTGAGGTGGGGGTGGGGGTGACTCTGCAAAATGGGGGTGTCCTCCGATCTCTGGCG 967  
969 AGGTGAGGAACAGACAGAGGGGGGTCTAGATTTCTGAGGGGGTGGTGGATATTTGGCAAG 1028  
968 AGGTGAGGAACAGACAGAGGGGGGTCTAGATTTCTGAGGGGGTGGTGGATATTTGGCAAG 1027  
1029 CAGGAAACCTCTGGAGAGCTCTCATTTTCTCCATGGGGAAGACAGAGCATCTCTTCAGGAG 1088  
1028 CAGGAAACCTCTGGAGAGCTCTCATTTTCTCCATGGGGAAGACAGAGCATCTCTTCAGGAG 1087  
1089 AGACTCCAGGGCAAGAGGGGTGTCTTGGCTGTGCTTGAAGGGGAAACCTTGCCATATC 1148  
1088 AGACTCCAGGGCAAGAGGGGTGTCTTGGCTGTGCTTGAAGGGGAAACCTTGCCATATC 1147  
1149 CCCAGTCCAGTCCCTCAGCCTGTGCTGGCTTGCATCTCCTGACTGGATGTTCTCAGGCC 1208  
1148 CCCAGTCCAGTCCCTCAGCCTGTGCTGGCTTGCATCTCCTGACTGGATGTTCTCAGGCC 1207  
1209 CTTGTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGGATCTAATAAACCTCTTGGAGCA 1268  
1208 CTTGTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGGATCTAATAAACCTCTTGGAGCA 1267  
1269 CAAAAAATAAAAA 1281  
1268 CAAAAAATAAAAA 1280

RESULT 2  
AAC02135  
ID AAC02135 standard; cDNA; 465 BP.  
XX AAC02135;  
XX 06-OCT-2000 (first entry)  
XX Human secreted protein 5' EST, SEQ ID NO: 2133.  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping; ss.  
XX Homo sapiens.  
XX OS  
XX EP1033401-A2.  
XX 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-00200610.  
XX 26-FEB-1999; 99US-0122487P.  
XX (GIST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX P-PSDB; AAG02129.  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX

Claim 1; SEQ ID NO 2133; 71pp + Sequence Listing; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

Sequence 465 BP; 92 A; 117 C; 160 G; 90 T; 0 U; 6 Other;

Query Match 35.8%; Score 459.2; DB 3; Length 465;  
Best Local Similarity 98.5%; Pred. No. 6.2e-117;  
Matches 458; Conservative 3; Mismatches 4; Indels 0; Gaps 0

1 GATTTTAGGGGTACGGAGAGTGTCAGCTTCAGGCATCGGAGCGCTGGCGCCCATGG 306  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
1 GATTTTAGGGGTACGGAGAGTGTCAGCTTCAGGCATCGGAGCGCTGGCGCCCATGG 60

307 CCCCGCTGGGAGGCCGCCCGCTGGTA CTGCTTTCAGCGCAAGAGAAATCCGGGA 366  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
61 CCCCGCTGGGAGGCCGCCCGCTGGTA CTGCTTTCAGCGCAAGAGAGAAATCCGGGA 120

367 AGGACTTCGTACACGAGGCGCTGCAGAGCAGACTTGGAGCTGATGCTGTGCTGCTCC 426  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
121 AGGACTTCGTACACGAGGCGCTGCAGAGCAGACTTGGAGCTGATGCTGTGCTGCTCC 180

427 GGCTCTCTGTGCTCACTCAAGGAA CAGTATGCTCAGGAGCATGCTTGA ACTTCCAGAGAC 486  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
181 GGCTCTCTGTGCTCACTCAAGGAA CAGTATGCTCAGGAGCATGCTTGA ACTTCCAGAGAC 240

487 TCCTTGGACACCAGCACCTACAAGGAGGCCCTTTCGGAAGGACATGATCCGCTGGGGAGAGG 546  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
241 TCCTTGGACACCAGCACCTACAAGGAGGCCCTTTCGGAAGGACATGATCCGCTGGGGAGAGG 300

547 AGAAACGCCAGGCTGACCCAGGCTCTTTTTCGAGGAAGATTTGAGAGGGGATCTCCAGC 606  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
301 AGAAACGCCAGGCTGNNNCAGGCTCTTTTTCGAGGAAGATTTGAGAGGGGATCTCCAGC 360

607 CCATCTGCTGCTGAGTGACACGAGAGTGTCGACATCCAGTGGTTCGAGAGGCGCT 666  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
361 CCATCTGCTGCTGAGTGACACGAGAGTGTCGACATCCAGTGGTTCGAGAGGCT 420

667 ATGGGGCCGTGACGAGCAGCGTCCGCTTGAGCGTTGGAGCAGA 711  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
421 ATGGGGCCGTGACGAGCAGCGTCCGCTTGAGCGTTGGAGCAGA 465

RESULT 3  
ACH36439  
ID ACH36439 standard; cDNA; 475 BP.  
AC ACH36439;  
XT 13-OCT-2003 (first entry)  
CX Human endothelial cell cDNA #4572.  
DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
CX genome mapping; biodiversity; genetic disorder.  
CX Homo sapiens.  
CX US2003073623-A1.

PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Dzmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
WP	WPI; 2003-615964/59.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
CS	Claim 1; SEQ ID NO 23651; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?DocID=20030073623
XX	
SQ	Sequence 475 BP; 107 A; 115 C; 150 G; 96 T; 0 U; 7 Other;
	Query Match            31.1%; Score 398.6; DB 8; Length 475;
	Best Local Similarity 97.8%; Pred. No. 4e-100;
	Matches 404; Conservative     0; Mismatches     9; Indels     0; Gaps     0;
QY	869 CTCGACTTTAGTCACTAGGTTC TAGGAGTAGCTGGGCCCTGCCTGAGTGCGGGTGGG 928
DB	34 CTCGCGGAATTCGACACGAGGTTC TAGGAGTAGCTGGGCCCTGCCTGAGTGCGGGTGGG 93
QY	929 GCTGACTCTGC AAAATGCGGGGTGTCCC CGATCCTCTGGCCGAGGTGAGGACACACAGGGG 988
DB	94 GCTGACTCTGC AAAATGCGGGGTGTCCC CGATCCTCTGGCCGAGGTGAGGAAACACACAGGGG 153
QY	989 GGGTCTAGATTCTGAGGGGGTTCGGTGGATATTGGGCAAGGCAGGAAAACCTCTCTGAGACCT 1048
DB	154 GGGTCTAGATTCTGAGGGGGTTCGGTGGATATTGGGCAAGGCAGGAAAACCTCTCTGAGACCT 213
QY	1049 CATTTCTCCATGGGGAAGACACACGCATGTCTTTCAGGAGAGACTCCAAGGCCAAAGGAG 1108
DB	214 CATTTTCTCATGGGGAAGACACGCATGTCTTTCAGGAGAGACTCCAAGGCCAAAGGAG 273
QY	1109 GGTGTCTTGSCCTGTGCTTGAAGCGGAAACCCTGCCATATCCCCAGTGCACAGTCCCTCAG 1168
DB	274 GGTGTCTTGSCCTGTGCTTGAAGCGGAAACCCTGCCATATCCCCAGTGCACAGTCCCTCAG 333
QY	1169 CCTGTGGTGCCCTTGATCTCTGACTGGATGTTCTCAGCCCTTGTCTGTGGCAGAACCC 1228
DB	334 CCTGTGGTGCCCTTGATCTCTGACTGGATGTTCTCAGCCCTTGTCTGTGGCAGAACCC 393
QY	1229 AGAGCTCCCCAGTGTGGATACTATAAACCTCTTGGAGCACAAAAAATAAAAA 1281

Db	394	AGAGCTCCCGAGTGGGATGATCTATAAACCCTTCTGGAGCACAAAAA	446
RESULT 4			
ACH29232			
XX	ACH29232	standard; cDNA; 407 BP.	
XX	ACH29232;		
AC	ACH29232;		
DT	13-OCT-2003	(first entry)	
XX	Human adult spleen cDNA #251.		
DE	Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;		
XX	genome mapping; biodiversity; genetic disorder.		
KW	Homo sapiens.		
XX	US2003073623-A1.		
XX	17-APR-2003.		
XX	30-JUL-2001; 2001US-00918995.		
FF	30-JUL-2001; 2001US-00918995.		
XX	(DRMA/) DRMANAC R T.		
PA	(LABA/) LABAT I.		
PA	(STAC/) STACHE-CRAIN B.		
PA	(DICK/) DICKSON M C.		
PA	(JONE/) JONES L W.		
XX	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;		
PI	WPI; 2003-615964/58.		
DR	New polynucleotide sequences obtained from various cDNA libraries, useful		
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene		
PT	mapping, in the recombinant production of protein, or in generating		
PT	antisense DNA or RNA.		
XX	Claim 1; SEQ ID NO 16444; 44pp; English.		
PS	The invention relates to an isolated polynucleotide comprising any one of		
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was		
CC	determined by the technique of SBH (sequencing by hybridisation). Also		
CC	included is a purified polypeptide comprising a sequence corresponding to		
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences		
CC	are useful in diagnostics as expressed sequence tags (EST) for		
CC	identifying expressed genes or for physical mapping of the human genome,		
CC	in forensics in assessing biodiversity, or in identifying mutations		
CC	responsible for genetic disorders and other traits. The nucleotide		
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,		
CC	for chromosome and gene mapping, in the recombinant production of		
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide		
CC	is useful for generating antibodies specific for it. The present sequence		
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data		
CC	for this patent did not form part of the printed specification, but was		
CC	obtained in electronic format directly from USPTO at		
CC	seqdata.uspto.gov/sequence.html?DocID=20030073623		
XX	Sequence 407 BP; 70 A; 120 C; 139 G; 78 T; 0 U; 0 Other;		
SQ	Query Match 30.8%; Score 394.2; DB 8; Length 407;		
	Best Local Similarity 98.0%; Pred. No. 6.2e-99;		
	Matches 399; Conservative 0; Mismatches 8; Indels 0; Gaps 0;		
QY	129 ATGAGTGTAGGACCGTCCGGCGGTTCAGCCCTGTTTCGAGCTACGGCAGAGAGGGCC	188	
Db	1 ACGAGTGTAGGACCGTCCGGCGGTTCAGCCCTGTTTCGAGCTACGGCAGAGAGGGCC	60	
QY	189 GCGATTGTTCTTGTTCGCGCTCCGCTTAGTGGCGCGTCCATTCCGCGCGGTGCCGA	248	

Db	61	GCGATTGCTCCTTGCTGCGCGCTCCGCTTAGTGGCGCGCTCCATTCCGCGCGGTCCCGA	120
QY	249	TTTTAGGGGTAGGAGAAAGTGTCTAGCTTCAGGCATCGCGAGCGTGGCGGCCCATGSCC	308
Db	121	TTTTAGGGGTAGGAGAAAGTGTCTAGCTTCAGGCATCGCGAGCGTGGCGGCCCATGSCC	180
QY	309	CGCTGGGAGCGCCCCCGCGCTGCTACTGTCTGTTTTCAGCGCAAGAGAAATCCGGGAAG	368
Db	181	CGCTGGGAGCGCCCCCGCGCTGCTACTGTCTGTTTTCAGCGCAAGAGAAATCCGGGAAG	240
QY	369	GACTTCGTGACCGAGGCGCTCGAGAGCAGACTTGGAGCTGATGTCTGCTGCTCTCCGG	428
Db	241	GACTTCGTGACCGAGGCGCTCGAGAGCAGACTTGGAGCTGATGTCTGCTGCTCTCCGG	300
QY	429	CTCTCTGTGTCACCTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAATTCAGAGACTC	488
Db	301	CTCTCTGTGTCACCTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAATTCAGAGACTC	360
QY	489	CTGGACACCGACCTTCAAGAGGCGCTTCGGAAGGACATGATCCG	535
Db	361	CTGGACACCGACCTTCAAGAGGCGCTTCGGAAGGACATGATCCG	407
RESULT 5			
AAT23873			
ID	AAT23873	standard; cDNA to mRNA; 337 BP.	
XX	AAT23873;		
AC	AAT23873;		
XX	16-AUG-1996	(first entry)	
DT	Human gene signature HUMGS05801.		
DE	Gene signature; messenger RNA; mRNA; relative abundance; frequency;		
XX	human; cloning; mapping; non-biased library; diagnosis; detection;		
KW	cell typing; abnormal cell function; ss.		
XX	Homo sapiens.		
OS	WO9514772-A1.		
XX	01-JUN-1995.		
XX	11-NOV-1994; 94WO-JP001916.		
XX	12-NOV-1993; 93JP-00355504.		
PR	(MATS/) MATSUBARA K.		
XX	(OKUB/) OKUBO K.		
PA	Matsubara K, Okubo K;		
XX	WPI; 1995-206931/27.		
DR	Single-stranded DNA for identifying gene signatures - isolated from 3'-		
XX	directed human cDNA library that reflects relative abundance of corresp.		
PT	mRNA in specific human tissues.		
PS	Claim 1; Page 1475; 2245pp; Japanese.		
XX	A single-stranded DNA (or its complementary strand or the corresp. double		
CC	-stranded DNA) which comprises one of the 7037 "GS" sequences given in		
CC	AAT19001-T26837 and which is able to hybridise to part of human genomic		
CC	DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were		
CC	obtained from 3'-directed cDNA libraries prepared from various human		
CC	tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using		
CC	poly(T) as the sole primer. Since the 3'-untranslated sequence is unique		
CC	to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise		
CC	with specific mRNAs. Each library is constructed so as to reflect		
CC	accurately the relative abundance of different mRNAs in the particular		
CC	tissue from which it was derived. The appearance frequency of a given GS		
CC	in a cDNA library can be determined (esp. using primers and probes		
CC	derived from the GS sequences) as a means of diagnosing abnormal cell		

```

CC function or for recognising different cell types
CC Sequence 337 BP; 82 A; 76 C; 94 G; 72 T; 0 U; 13 Other;
Query Match 22.8%; Score 292.6; DB 2; Length 337;
Best Local Similarity 92.9%; Pred. No. 8.4e-71;
Matches 301; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
NY 958 GATCCTGGCCGAGTGGAGAACAGACAGAGGGGGGCTAGATCTGAGGGGGTGGTGGAT 1017
b 1 GATCCTGGCCGAGTGGAGAACAGACAGAGGGGGGCTAGATCTGAGGGGGTGGTGGAT 60
y 1018 ATTGGGCAAGCAGGAAACCTCTGGAGACCTATTTCTCCATGGGGAAGACAGCCATGC 1077
b 61 ATTGGGCAAGCAGGAAACCTCTGGAGACCTATTTCTCCATGGGGAAGACAGCCATGC 120
y 1078 TCTTCAGGAGGAGACTCCAAAGGCAAGAGGAGGTGTCTTGGCTGTCTTGAAGGCGAAAC 1137
b 121 TCTTCAGGAGGAGACTCCAAAGGCAAGAGGAGGTGTCTTGGCTGTCTTGAAGGCGAAAC 180
y 1138 CTTGCCATATCCCGAGTCCAGTCCCTCAGCTGTGTGGTGGCTTCATCCTGACTGGAT 1197
b 181 CTTGCCATATCCCGAGTCCAGTCCCTCAGCTGTGTGGTGGCTTCATCCTGACTGGAT 240
y 1198 GTTCTCAGCCCTTGTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGGATCTAATAAAC 1257
b 241 GTTCTCAGCCCTTGTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGGATCTAATAAAC 300
y 1258 CTCTTGAGGACACAAAAA 1281
b 301 CCTNTTGGAGCNCNCAATANAA 324

RESULT 6
ABL20639/C
ID AAL20639 standard; cDNA; 502 BP.
CC
AC AAL20639;
CC
CT 07-DEC-2001 (first entry)
DE Human breast cancer expressed polynucleotide 13096.
CC
KW Human; breast cancer; cell marker; cytostatic; ss.
CC
DS Homo sapiens.
CC
SN WO200151628-A2.
CC
PD 19-JUL-2001.
CC
PF 10-JAN-2001; 2001WO-US000798.
CC
PR 14-JAN-2000; 2000US-0176077P.
CC
PR 14-MAR-2000; 2000US-0189167P.
CC
PR 24-MAR-2000; 2000US-0192099P.
CC
PR 29-MAR-2000; 2000US-0193480P.
CC
PR 15-MAY-2000; 2000US-0205230P.
CC
PR 09-JUN-2000; 2000US-0211315P.
CC
PR 25-JUL-2000; 2000US-0220534P.
CC
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
CC
PI Lillie J, Xu Y, Wang Y, Steinmann K;
CC
WPI; 2001-451856/48.
CC
PT New peptide useful as a marker for the diagnosis of breast cancer.
CC
DS Claim 1; Page 2324; 3695pp; English.
CC
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is

```

```

CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
CC
SQ Sequence 502 BP; 131 A; 148 C; 134 G; 89 T; 0 U; 0 Other;
Query Match 9.8%; Score 125; DB 4; Length 502;
Best Local Similarity 96.2%; Pred. No. 2.9e-24;
Matches 128; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 337 TGCTGTTCAGCGGCAAGAGGAATCCGGAAGGACTTCGTGACCGAGGCGCTGCAGAGCA 396
b 495 TGCTGTTCAGCGGCAAGAGGAATCCGGAAGGACTTCGTGACCGAGGCGCTGCAGAGCA 436
QY 397 GACTTGGAGCTGATGTCTGTGCTCTCTCCGCGCTCTCTGCTCAAGGAACAGTATG 456
b 435 GACTTGGAGCTGATGTCTGTGCTCTCTCCGCGCTCTCTGCTCAAGGAACAGTATG 376
QY 457 CTCAGGAGCATGG 469
b 375 CTCAGGAGCATGG 363

RESULT 7
ABL30119
ID ABL30119 standard; DNA; 570 BP.
CC
AC ABL30119;
CC
CT 26-MAR-2002 (first entry)
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 41830.
CC
KW Drosophila; developmental biology; cell signalling; insecticide;
CC pharmaceutical; gene; ds.
CC
OS Drosophila melanogaster.
CC
PN WO200171042-A2.
CC
PD 27-SEP-2001.
CC
PF 23-MAR-2001; 2001WO-US009231.
CC
PR 23-MAR-2000; 2000US-0191637P.
CC
PR 11-JUL-2000; 2000US-00614150.
CC
PA (PEKE ) PE CORP NY.
CC
PI Venter JC, Adams M, Li PWD, Myers EW;
CC
WPI; 2001-656860/75.
CC
PT New isolated nucleic acid detection reagent for detecting 1000 or more
CC genes from Drosophila and for elucidating cell signaling and cell-cell
CC interactions.
CC
PS Claim 1; SEQ ID NO 41830; 21pp + Sequence Listing; English.
CC
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

```

XX  
SQ Sequence 570 BP; 127 A; 148 C; 184 G; 111 T; 0 U; 0 Other;  
Query Match 6.4%; Score 82; DB 4; Length 570;  
Best Local Similarity 51.9%; Pred. No. 2.6e-12;  
Matches 268; Conservative 0; Mismatches 230; Indels 18; Gaps 3;  
QY 337 TGCTGTTCCAGCGGCAAGAGAAATCCCGGAAGGACTTCGTACCGAGGCGCTGCAGAGCA 396  
Db 14 TACTCATCAGCGGCAAGAGAAAGTGGCGGAGAGATACATATCCGAGAGCTGCAGCGGA 73  
QY 397 GACTTGGAGCTGATGTCTGTCTCTCTCCGGCTCTCTGTCTCCACTCAAGGAACAGATG 456  
Db 74 GATTGGGCTCCCGCTCGTGA---TCGTTCAATCTCAGAGCCCATTAAGTCCGAATGGG 130  
QY 457 CTCAGGAGCATGGTGAACCTTCCAGAGACTCTCTGGACACAGCACCTCAAGAGGSCCT 516  
Db 131 CTCGCAAGCTCAGTTGACCTGACCGCTCTCTCCGTGATGGACCTACAGAGAGT 190  
QY 517 TTCGGAAGGACATGCTCCGCTGGGAGAGAGAAACCCAGGCTGACCCAGGCTCTTTT 576  
Db 191 ACCGCGCGACATGATTGTGTGAGAGACGAGGTGCGGGCCAGACTACGGCTACTTCT 250  
QY 577 GCAGGAGATTGTGAGG-----GCATCTCCAGCCCATCTGGCTGTGAGTG 624  
Db 251 GCCGCTGGCATGAGAGAGCGCTGATGCCAGCAGACGCCCTACATTTCTGTGAGCG 310  
QY 625 ACACACGAGAGATGCTGACATCCAGTGGTTTCGGGAGGCTATGSGGCCG---TGACGC 681  
Db 311 ATGTGCGCGCAAGAACGACATCAGGTGGTTTCGGGAGACTTACGGCCCGAACGAGTCA 370  
QY 682 AGAGGTCGGGTTGTAGCTGTGAGCAGACGACCCAGCAGCGGGGCTGGTGTTCAGCG 741  
Db 371 TCACCTTCGGTCTAACATCCCGTCCAGAAACCGGTTCTGCACGGGGATGACCTTTACCG 430  
QY 742 CAGGGCTGCGAGTCTGAGTACAGATGTGCCCTGGACAACTTCGGGAGCTTTGACTGGG 801  
Db 431 CAGGATAGACGATGTTCCCTCCGAGTCCGATCTGGATGATCTGCGCAGCGCTTCAGC 490  
QY 802 TCATCGAAGACATGGAGTTGAAACAGCGCCCTGGAGG 837  
Db 491 TGGTGTGGCCAAACAGCAGGAGAACTCGACCCAGGAGG 526

RESULT 8  
AAL11749/c  
ID AAL11749 standard; cDNA; 568 BP.  
AC AAL11749;  
DT 07-DEC-2001 (first entry)  
DE Human breast cancer expressed polynucleotide 4206.  
XX Human; breast cancer; cell marker; cytostatic; ss.  
XX Homo sapiens.  
XX WO200151628-A2.  
XX 19-JUL-2001.  
XX 10-JAN-2001; 2001WO-US000798.  
XX 14-JAN-2000; 2000US-0176077P.  
XX 14-MAR-2000; 2000US-0189167P.  
XX 24-MAR-2000; 2000US-0192099P.  
XX 29-MAR-2000; 2000US-0193480P.  
XX 15-MAY-2000; 2000US-0205230P.  
XX 09-JUN-2000; 2000US-0211315P.  
XX 25-JUL-2000; 2000US-0220534P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX  
PI Lillie J, Xu Y, Wang Y, Steinmann X;  
XX WPI; 2001-451856/48.  
XX New peptide useful as a marker for the diagnosis of breast cancer.  
XX Claim 1; Page 767; 3695pp; English.  
XX The invention relates to human breast cancer expressed polynucleotides  
XX (AAL07544-AAL26789) and methods of assessing whether a patient is  
XX afflicted with breast cancer by examining the correlation between the  
XX expression of certain markers and the cancerous state of breast cells.  
XX The polynucleotides and encoded polypeptides are potential markers for  
XX detecting, diagnosing, monitoring, characterizing treating and  
XX potentially preventing breast cancer. The polynucleotides and encoded  
XX polypeptides are also useful for isolating compounds with cytostatic  
XX activity  
XX SQ Sequence 568 BP; 142 A; 170 C; 152 G; 100 T; 0 U; 4 Other;  
Query Match 6.0%; Score 76.4; DB 4; Length 568;  
Best Local Similarity 76.8%; Pred. No. 9.2e-11;  
Matches 119; Conservative 0; Mismatches 32; Indels 4; Gaps 2;  
QY 319 GCGCCCGCGGCTGCTACTGCTGTTCACGCGCAGAGAAATCCGGAA--GGACTTCGT 376  
Db 549 GCGGCGCGGAGGTACTTCTGTTTTCAGCGGCGAAGAGAAATCCGGAAAGACTTTCGT 490  
QY 377 GACCGA--GGCGCTCAGAGCAGACTTGGAGCTGATGTCTGTCTGTCTCCGGCTCTCT 434  
Db 489 GACCGAGCGGCTGCAAGAGCAGACTTGGAGCTGATGTCTGTCTGTCTCCGGCTCTCT 430  
QY 435 GGTCCACTCAGGAACAGATGCTCAGGAGCATGG 469  
Db 429 GGTCCACTCAGGAACAGATGCTCAGGATGGTGG 395  
RESULT 9  
ABL30118/c  
ID ABL30118 standard; DNA; 2673 BP.  
XX ABL30118;  
XX 26-MAR-2002 (first entry)  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 41827.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ds.  
XX Drosophila melanogaster.  
XX WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX Claim 1; SEQ ID NO 41827; 21pp + Sequence Listing; English.



XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
XX ABBS72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 2673 BP; 610 A; 738 C; 638 G; 697 T; 0 U; 0 Other;  
XX  
XX Query Match 4.7%; Score 60.8; DB 4; Length 2673;  
XX Best Local Similarity 50.0%; Pred. No. 3.7e-06;  
XX Matches 217; Conservative 0; Mismatches 202; Indels 15; Gaps 2;  
XX  
XX 419 TGCTCTCCGGCTCTCTGCTCACTCAAGAACAGTATGCTCAGGACATGGCTTGAACCT 478  
XX 1478 TATCGTTTCAATCTCAGAGCCATTAAAGTCGGAATGGGCTCGCAAGCTCAGTTGACCT 1419  
XX 479 CCAGAGACTCTGTGACACACAGACCTCAAGAGGAGGCTTTCCGAAGAGACATATCGCTG 538  
XX 1418 GGACGCTCTCTCGGTGATGACCTTACAGAGAGAGTACCGGCGGACATGATTGTGTG 1359  
XX 539 GGGAGAGAGAAACGCCAGGCTGACCGCTCTCTTTTTCAGGAAGATTGTGGAGG---- 594  
XX 1358 GAGCGACGAGGTGCGGGCCAGGACTACGGCTACTTCTGCCGCTGGCCATGGAGGAGGC 1299  
XX 595 -----GCATCTCCAGCCCATCTGGCTGGTGGTGAACACAGGAGTGTCTGACAT 646  
XX 1298 GCTGAGTCCGACGACAGCGCCCTACATCTTGCTGCGGATGCTGCGCGCAGAACGACAT 1239  
XX 647 CCAGTGTGTTTCGGAGGCTATGGGCGGCTGACGACAGCGTCCGCGTGTAGCGTTTGA 706  
XX 1238 CAGGTGTTTCGGAGACTTACGGGCGGAAACGAGTATCACCTTGGCTTAAACATCCCG 1179  
XX 707 GC---AGAGCGGACAGCAGCGGGCTGGGTGTTTCAGCGAGGGTGGAGTGTCTGAGTC 763  
XX 1178 TCAGAAACCGGGTCTGACGGGATGGACCTTTACCGCAGGAATAGACGATGTTCCCTC 1119  
XX 764 AGAATGTGGCTGGACAACTTCGGGGACTTTGACTGGGTATCGAAGCCATGGAGTTGA 823  
XX 1118 CGAGTGGATCTGGATGATCTGGCGGACGGCTTCGACGTGGTGTGGCCAAACGACGAGGA 1059  
XX 824 ACAGCGCTGGAGG 837  
XX 1058 ACTCGACGAGG 1045  
XX  
XX RESULT 10  
XX ABL30184/C  
XX ID ABL30184 standard; DNA; 3406 BP.  
XX AC ABL30184;  
XX AC ABL30184;  
XX XX  
XX 26-MAR-2002 (first entry)  
XX DT  
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 42025.  
XX XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmacological; gene; ds.  
XX OS Drosophila melanogaster.  
XX FN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX PF  
XX 23-MAR-2000; 2000US-0191637P.  
XX PR

PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 42025; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
XX ABBS72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 3406 BP; 965 A; 833 C; 729 G; 879 T; 0 U; 0 Other;  
XX  
XX Query Match 4.7%; Score 60.8; DB 4; Length 3406;  
XX Best Local Similarity 50.0%; Pred. No. 4.1e-06;  
XX Matches 217; Conservative 0; Mismatches 202; Indels 15; Gaps 2;  
XX  
XX 419 TGCTCTCCGGCTCTCTGCTCACTCAAGAACAGTATGCTCAGGACATGGCTTGAACCT 478  
XX 3209 TATCGTTTCAATCTCAGAGCCATTAAAGTCGGAATGGGCTGCGAGCTGCGAGTTGGACCT 3150  
XX 479 CCAGAGACTCTGTGACACACGACCTACAGAGGCGCTTTTCGGAAGGACATGATCCGCTG 538  
XX 3149 GGACGCTCTCTCGGTGATGAGACCTTACAGGAGAGTACCGGCGGACATGATTGTGTG 3090  
XX 539 GGGAGAGAGAAACGCCAGGCTGACCGGCTCTTTTTCAGGAAGATTGTGGAGG---- 594  
XX 3089 GAGCGACGAGGTGCGGGCCAGGACTACGGCTACTTCTGCCGCTGGCCATGGAGGAGGC 3030  
XX 595 -----GCATCTCCAGCCCATCTGGCTGGTGGTGAACACAGGAGTGTCTGACAT 646  
XX 3029 GCTGAGTCCGACGACAGCGCCCTACATCTCTGTCAGCGATGTGCGGCGCAAGAACGACAT 2970  
XX 647 CCAGTGTGTTTCGGAGGCTATGGGCGGCTGACGACAGCGTCCGCGTGTAGCGTTTGA 706  
XX 2969 CAGGTGGTTCGGGAGACTTACGGGCGGAAACGAGTATACCTTGGCTTAAACATCCCG 2910  
XX 707 GC---AGAGCGGACAGCAGCGGGCTGGGTGTTTCACGCGGAGGTGGACGATGCTGAGTC 763  
XX 2909 TCCAGAAACCGGGTCTGACCGGGATGGACCTTTTACCGCAGGAATAGACGATGTTCCCTC 2850  
XX 764 AGAATGTGGCTGGACAACTTCGGGGACTTTGACTGGGTATCGAAGCCATGGAGTTGA 823  
XX 2849 CGAGTGGATCTGGATGATCTGGCGGACGGCTTCGACGTGGTGTGGCCAAACGACGAGGA 2790  
XX 824 ACAGCGCTGGAGG 837  
XX 2789 ACTCGACGAGG 2776  
XX  
XX RESULT 11  
XX ABL35477  
XX ID ABL35477 standard; DNA; 60 BP.  
XX AC ABL35477;  
XX AC ABL35477;  
XX XX  
XX 15-JUL-2002 (first entry)  
XX DT  
XX Human spliced transcript detection oligonucleotide SEQ ID NO:8225.  
XX DE



KX Human; mouse; rat; splice transcript; detection; RNA transcript;  
 KW splice variant; transcriptome; oligonucleotide library; ss.  
 XX Homo sapiens.  
 DS WO200210449-A2.  
 FN 07-FEB-2002.  
 PD 20-JUL-2001; 2001WO-IB001903.  
 PF 28-JUL-2000; 2000US-0221607P.  
 PR 02-MAY-2001; 2001US-0287724P.  
 XX (COMP-) COMPUEN INC.  
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 FI WPI; 2002-257383/30.  
 DR New oligonucleotide libraries comprising oligonucleotides which  
 XX selectively hybridize to mRNAs transcribed from a transcription unit of a  
 XX genome, useful for detecting tissue-, pathology-, and developmental-  
 XX specific genes.  
 PS Example 1; SEQ ID NO 8225; 47pp; English.  
 CC The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
 CC )transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises several  
 CC oligonucleotides, each capable of hybridizing selectively to a set of  
 CC messenger RNAs transcribed from a given transcription unit of the genome,  
 CC which encodes one or more messenger RNA splice variants. The  
 CC oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or  
 CC quantitatively characterizing the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialized mini  
 CC libraries to detect transcripts of a sub-transcriptome under a particular  
 CC biological or pathological state, and so allowing the detection of tissue  
 CC - and pathology-specific genes such as those genes only expressed in  
 CC specific tissue under a specific pathological condition; to detect  
 CC developmental specific genes; and to detect RNA transcripts and splice  
 CC variants of a transcriptome of a patient suffering from a particular  
 CC disorder. ABN7253 to ABN9589 represent oligonucleotide sequences from  
 CC rate, humans and mice, which are used in the exemplification of the  
 CC present invention. N.B. The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 60 BP; 13 A; 17 C; 15 G; 15 T; 0 U; 0 Other;  
 SQ Query Match 4.7%; Score 60; DB 6; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1194 GGATGTTCTACACCCCTGTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATCTAAAT 1253  
 DB 1 GGATGTTCTACACCCCTGTTCTGGGCAAGAACCCAGAGCTCCCGAGTGTGATCTAAAT 60  
 RESULT 12  
 ADA71938  
 ID ADA71938 standard; DNA; 2000 BP.  
 XX ADA71938;  
 AC ADA71938;  
 XX 20-NOV-2003 (first entry)  
 DT Rice gene, SEQ ID 5263.  
 DE Rice gene, SEQ ID 5263.  
 XX

KW Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.  
 XX Oryza sativa.  
 XX WO2003000898-A1.  
 XX 03-JAN-2003.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX WPI; 2003-175290/17.  
 DR Identifying at least one gene involved in plant resistance or response to  
 XX pathogenic infection for conferring resistance or tolerance to a plant to  
 XX bacterial, fungal or viral infection by determining or detecting plant  
 XX gene expression.  
 PS Claim 27; SEQ ID NO 5263; 899pp; English.  
 XX The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
 SQ Query Match 4.6%; Score 59.4; DB 7; Length 2000;  
 Best Local Similarity 9.1%; Pred. No. 8.1e-06;  
 Matches 66; Conservative 325; Mismatches 336; Indels 0; Gaps 0;  
 QY 286 CGAGCGTGGGGCCCATGGCCCGCTGGGAGGCGCCCGCGGTGTACTGTGTCA 345  
 DB 1 S2ACAGWKSCSSNRGRRMTYMGWMSCARMSRMRKMGMSKYRKCSCGKCKMTT 60  
 QY 346 GCGCAAGAGGAAATCCGGAGGACTTCGTGACGAGCGCTGCAGACGACACTGGAG 405  
 DB 61 RRSKMYSSASSAGRTGSKWSSGYSYSGMKKRYKRWGRGRGRGRSRMRMRWGR 120  
 QY 406 CTGATGTCGTCT 465  
 DB 121 RCARSGRMAGSGRMMGGKSRMSYMWVCYARGCGSKRKSXGSGWGTCTCRRGAGSGW 180  
 QY 466 ATGGCTTGAACTTCAGAGACTCTCTGGACACGACACCTCAAGAGAGGCTTTCCGAGG 525  
 DB 181 SGAATYKSGMSKRWMSGCGRSGGRRSAYSYRYGTSRKYGTYKMTYYSASRCRAYM 240  
 QY 526 ACATGATCTCGCTGGGAGAGAGAAACCGCAGGCTGACCCAGGCTTCTTTTCAGAGAGA 585  
 DB 241 TTSYSWACSSYTWCRSKRRSMWMMKMRWBSRYGYWSYKMMWMTAYKKSYYSRWCY 300  
 QY 586 TTGTGAGAGGCAATCTCCAGCCCATCTGGCTGTGAGTGACACACGAGAGTGTCTGACA 645  
 DB 301 MYRGGGWRGATYRWGRGYMRMMYKMYWTRGYKMGKRWGMMWMMRSMRSMRSMRSMR 360  
 QY 646 TCCAGTGGTTCGGGAGGCGCTATGGGCGCGTGCAGCAGACGCTCCCGCTTGTAGCGTTGG 705  
 DB 361 YMEWMWMTTRRRWAKSSRTSRKRWKMRKRWKMRKRWKMRKRWKMRKRWKMRKRW 420  
 QY 706 AGCAGAGCCGACAGCAGCGCGGCTGGTGTTCACGCCGGGCTGGACGATGCTGATCAG 765



PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI: 2003-029926/02.  
 DR P-PDS; AB033973.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 14; SEQ ID NO 25713; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2100 BP; 364 A; 732 C; 719 G; 285 T; 0 U; 0 Other;  
 Query Match 4.0%; Score 51; DB 7; Length 2100;  
 Best Local Similarity 52.5%; Pred. No. 0.0018;  
 Matches 134; Conservative 0; Mismatches 120; Indels 1; Gaps 1;  
 QY 121 GTTGGACATGAGCTGAGGCGACGGTCCGGCGGCTCAGCCCTGTCCAGCTACGGCGA 180  
 DB 1620 GTTCGCAAGGAGCGCGCGGAGAGAGCCCTGAGGCGCGGAGGACGCGATGGCCCGCA 1679  
 QY 181 GGAGGGCGCGATGTTCTTGTCCGCTCCGCTTAGTGGCGCGGTTCATTCGCGCGG 240  
 DB 1680 GTTCCCGCGCGCGGAGCTTCGCTTGGAGCGCTTGTCTCATGCGGATCA-TGAGCGCGG 1738  
 QY 241 TGTCCCGATTTTGGGGTAGGAGAGTGTCTAGCTTCAGGATCCGAGGGGTGGCGGCC 300  
 DB 1739 TGGCCCGCGAATCGACTACGCGAGCGTGTGGCCCTCTACACCGCGATCGGTGAGGCGC 1798  
 QY 301 CCATGCGCGCGTGGAGGCGCGCGCGCTGTCTGTCTGCTGAGCGGAGGAAAT 360  
 DB 1799 ATGTGTCGCGCGCGGACAGTGGTGTGAGCGGCTGTGTGCGCGGATCGACACAGG 1858  
 QY 361 CCGGGAAGGACTTCG 375  
 DB 1859 CCGAGGAGGACCTCG 1873  
 RESULT 15  
 AAC93822/c

ID AAC93822 standard; cDNA; 500 BP.  
 XX  
 AC AAC93822;  
 XX  
 DT 19-FEB-2001 (first entry)  
 XX  
 DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:317.  
 XX  
 KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;  
 KW flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;  
 KW detection; ss.  
 XX  
 OS Ctenocephalides felis.  
 XX  
 PN WO2000061621-A2.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 07-APR-2000; 2000WO-US009437.  
 XX  
 PR 09-APR-1999; 99US-0128704P.  
 XX  
 PA (HESK-) HESKA CORP.  
 XX  
 PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;  
 XX  
 DR WPI; 2000-656323/63.  
 XX  
 XX Flea Malpighian tubule and head and nerve cord tissue derived nucleic  
 PT acids useful for the prevention, diagnosis and treatment of flea  
 PT infestations.  
 XX  
 PT Claim 26; Page 350; 964pp; English.  
 PS  
 CC The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic  
 CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue  
 CC or head and nerve cord (HNC) tissue. The invention also relates to the  
 CC encoded proteins. The invention additionally encompasses expression  
 CC constructs, recombinant viruses and recombinant cells comprising the  
 CC nucleic acids of the invention, recombinant production of the proteins,  
 CC antibodies against the proteins, a method of identifying inhibitors of  
 CC the proteins, and compositions comprising the inhibitors for  
 CC administration to an animal. The nucleic acids, and the proteins they  
 CC encode may be used in the prevention, treatment and diagnosis of diseases  
 CC associated with flea infestations. For example, the nucleic acids may be  
 CC used to produce an HMT or HNC protein according to standard recombinant  
 CC DNA methodology by inserting the nucleic acids into a host cell and  
 CC culturing the cell to express the protein. The HMT and HNC nucleic acids  
 CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect  
 CC and quantitate the presence of cat flea or other homologous nucleic acid  
 CC sequences in samples. They may also be used to study the expression and  
 CC function of the proteins and their role in metabolism. The HMT and HNC  
 CC proteins may be used as antigens in the production of specific  
 CC antibodies, and in assays to identify modulators (agonists and  
 CC antagonists) of HMT and/or HNC protein expression and activity. The anti-  
 CC HMT/HNC protein antibodies and antagonists may also be used to  
 CC downregulate protein expression and activity. The antibodies may also be  
 CC used as diagnostic agents for detecting the presence of flea polypeptides  
 CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The  
 CC present sequence represents a cat flea HMT cDNA of the invention  
 XX  
 SQ Sequence 500 BP; 157 A; 81 C; 43 G; 200 T; 0 U; 19 Other;  
 Query Match 3.8%; Score 48.8; DB 3; Length 500;  
 Best Local Similarity 48.4%; Pred. No. 0.0039;  
 Matches 113; Conservative 1; Mismatches 126; Indels 0; Gaps 0;  
 QY 566 AGGCTTCTTTTGGAGGAATTTGGAGGCAATCTCCAGCCCATCTGGCTGTGAGTGA 625  
 DB 441 ATGGTTATTTTGTAAAGGCGTTGCCAAGCAATTTCAAACCTATTTCATAGTAGTGA 382  
 QY 626 CACACGAGAGTGTCTGACATCCAGTGGTTTCGGGAGCGCTATGGGCGCCGTGACGAGAC 685

Db 381 TATCCGAGAACACAGATATATAATGGTTATTGAAAAATATGCTAAAGTTATPAC 322  
 Oy 586 GGTCGGCGTTGTAGCGTTGAGACAGAGCCGACAGAGGGGGCTGGGTGTTCAAGCCAGG 745  
 Db 321 NATAGAAATTATGCACTAATGCAATTCGAAATATGAGGTTGAAATTTACNGAAG 262  
 Oy 746 GGTGACGATGCTGAGTCAAGATGTGGCTGACCACTTCGGGACTTTGACTGGGTCAAT 805  
 Db 261 TGTGATGATGTGCTGCTGATGATTTATATATATGAAACTGGGATTTGKCT 202  
 Oy 806 CGAGAA 811  
 Db 201 TGTTRA 196

Search completed: March 10, 2004, 10:59:01  
 Job time : 315.77 secs









785 ACTTCGGGAGCTTTGACTGGGTTCATCGAAGAACCATGGAGTTGAACAGCGCCTCGAGGAGC 844  
 841 AGTTGGAGAACCTCATAGAAATTTATCCGCTCCAGACTTTTGTCTAGTCTAGGTTCTAGGAGTG 900  
 845 AGTTGGAGAACCTCATAGAAATTTATCCGCTCCAGACTTTTGTCTAGTCTAGGTTCTAGGAGTG 904  
 901 AGCTGGGGCCTGCTGAGGTGGGGGTGGGGCTGACTCTGCAAAATGGGGGTGTCCTCCCGAT 960  
 905 AGCTGGGGCCTGCTGAGGTGGGGGTGGGGCTGACTCTGCAAAATGGGGGTGTCCTCCCGAT 964  
 961 CCTGGCCGAGTGAAGAACACAGACAGGGGGGTCTAGATTCTGAGGGGGTGTGGTGGATATT 1020  
 965 CTTGGCCGAGTGAAGAACACAGACAGGGGGGTCTAGATTCTGAGGGGGTGTGGTGGATATT 1024  
 1021 GGGCAAGGCAAGAACTCTTGAGAGACTCTATTTCTCCATGGGGAACAGACCATGCTCT 1080  
 1025 GGGCAAGGCAAGAACTCTTGAGAGACTCTATTTCTCCATGGGGAACAGACCATGCTCT 1084  
 1081 TCAGGAGGAGACTCAAGGGCAAGAGAGGGTCTTGGCTGTGCTTGAAGCGCAAAACCTT 1140  
 1085 TCAGGAGGAGACTCAAGGGCAAGAGAGGGTCTTGGCTGTGCTTGAAGCGCAAAACCTT 1144  
 1141 GCCATATCCCGAGTCCAGTCCCTCAGCCCTGCTGGTGGCTTGCATCTCCTGACGTGGATTT 1200  
 1145 GCCATATCCCGAGTCCAGTCCCTCAGCCCTGCTGGTGGCTTGCATCTCCTGACGTGGATTT 1204  
 1201 CTCAGCCCTTGTCTGGGCAAGAACCCAGAGCTCCCGAGTGGTGGATCTAATAAACCTC 1260  
 1205 CTCAGCCCTTGTCTGGGCAAGAACCCAGAGCTCCCGAGTGGTGGATCTAATAAACCTC 1264  
 1261 TTGGGACACAAAAA 1281  
 1265 TTGGGACACACTCTCAAAACA 1285

RESULT 2  
 BC007694  
 LOCUS Homo sapiens phosphomevalonate kinase, mRNA (cDNA clone MGC:11362 IMAGE:3954909), complete cds.  
 DEFINITION  
 BC007694  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1225)  
 Strausberg, R.D., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.G., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932  
 2 (bases 1 to 1225)  
 Strausberg, R.

TITLE  
 JOURNAL  
 REMARK  
 COMMENT  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 On Aug 19, 2003 this sequence version replaced gi:14043402.  
 Contact: MGC help desk  
 Email: cgapbs-re@mail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
 Web site: http://www.nisc.nih.gov/  
 Contact: nisc mgc@nigri.nih.gov  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirip, S., Thomas, P.O., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
 Clon distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IMAGE Plate: 15 Row: k Column: 14.  
 Location/Qualifiers  
 1..1225  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:11362 IMAGE:3954909"  
 /tissue\_type="Skin, melanotic melanoma."  
 /clone\_lib="NIH\_MGC\_20"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..1225  
 /gene="PMVK"  
 /note="synonym: PMKASE, PMK, HUMPMKI, PMKA"  
 /db\_xref="LOCUSID:10654"  
 /db\_xref="MIM:607622"  
 237..815  
 /codon\_start=1  
 /product="phosphomevalonate kinase"  
 /protein\_id="AAH07694.1"  
 /db\_xref="GI:14043403"  
 /db\_xref="LOCUSID:10654"  
 /translation="MAPCGAPRLVLPFGKSKGKDFVTEALQSRGLGADYCAVRLS  
 GELKEQYAOEHLNFORLDTSTYKEAPRKQMRKGEKQADPFQFKRIVEGISQ  
 IWLVSDFTRVSDIQWFRAYGAVTQTVRVVALESQRQGRQWVFTPGVDDASECGLDN  
 FGDFDWIENHGVQRLEQLELIEFIRSL"  
 276..623  
 /note="P-mevalon kinase; Region: Phosphomevalonate kinase.  
 Phosphomevalonate kinase (EC:2.7.4.2) catalyzes the  
 phosphorylation of 5-phosphomevalonate into  
 5-diphosphomevalonate, an essential step in isoprenoid  
 biosynthesis via the mevalonate pathway. This family  
 represents the animal type of the enzyme. The other is the  
 ERG8 type, found in plants and fungi, and some bacteria  
 (see pfam00288)"  
 /db\_xref="CDD:pfam04275"

FEATURES  
 source  
 gene  
 CDS  
 misc\_feature  
 ORIGIN  
 Query Match 94.8%; Score 1215; DB 9; Length 1225;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-290;  
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 67 CTGTGAGGAGACGGCTTTCTCTCCGCGTGTCTCACCTTTCTCCCATATCTGTTGG 126  
 Db 1 CTGTGAGGAGACGGCTTTCTCTCCGCGTGTCTCACCTTTCTCCCATATCTGTTGG 60



```

3b 139 CTGATGTCGTGTCCTCCGCGCTCTCTGGTCCACTCAAGGAACAGTATGCTCAGGAC 198
2y 466 ATGGCTTGAACCTTCCAGAGACTCCTGGACACACAGCACTACAAAGAGGCGCTTTTCGGAAG 525
3b 199 ATGGCTTGAACCTTCCAGAGACTCCTGGACACACAGCACTACAAAGAGGCGCTTTTCGGAAG 258
2y 526 ACATGATCCGCTGGGAGAGAGAAAGCCAGAGTGCACCGAGCTTCTTTTGAGGAAGA 585
3b 259 ACATGATCCGCTGGGAGAGAGAAAGCCAGAGTGCACCGAGCTTCTTTTGAGGAAGA 318
2y 586 TTGTGGAGGCGATCTCCAGCCCATCTGGCTGGTGTGAGTGACACACGAGAGTGTCTTGACA 645
3b 319 TTGTGGAGGCGATCTCCAGCCCATCTGGCTGGTGTGAGTGACACACGAGAGTGTCTTGACA 378
2y 646 TCAGTGGTTTGGGAGGCTATGGGCGCTGAGCGAGAGCGTCCGCTTGTAGCGTTGG 705
3b 379 TCAGTGGTTTGGGAGGCTATGGGCGCTGAGCGAGAGCGTCCGCTTGTAGCGTTGG 438
2y 706 AGCAGAGCCACACAGCGGGCTGGGTGTTTCAGCCAGGGGTGGAGTGTGAGTGTGAG 765
3b 439 AGCAGAGCCACACAGCGGGCTGGGTGTTTCAGCCAGGGGTGGAGTGTGAGTGTGAG 498
2y 766 AATGTGCTTGGACAACTTCGGGAGCTTTGACTGGGTGTCATCGAGAACCATGGAGTTGAAC 825
3b 499 AATGTGCTTGGACAACTTCGGGAGCTTTGACTGGGTGTCATCGAGAACCATGGAGTTGAAC 558
2y 826 AGCGCTTGGAGGACGTTGGAGAACCTGATAGATTTATCCGCTCCAGACTTTAGTTCAC 885
3b 559 AGCGCTTGGAGGACGTTGGAGAACCTGATAGATTTATCCGCTCCAGACTTTAGTTCAC 618
2y 886 TAGGTTCTAGAGTGTGAGTGGGCTCTGAGGTGGGGTGGGGTGTGACTCTGCAAAATG 945
3b 619 TAGGTTCTAGAGTGTGAGTGGGCTCTGAGGTGGGGTGGGGTGTGACTCTGCAAAATG 678
2y 946 GGGGTGTCCCGGATCTGCGCGAGGTGAGGAACAGACAGCGGGGTCTAGATTCGTGAGG 1005
3b 679 GGGGTGTCCCGGATCTGCGCGAGGTGAGGAACAGACAGCGGGGTCTAGATTCGTGAGG 738
2y 1006 GGGGTGTGAGTATTTGGGCAAGCAGGAACCTCTGGAGACCTCATTTTCTCATCGGGA 1065
3b 739 GGGGTGTGAGTATTTGGGCAAGCAGGAACCTCTGGAGACCTCATTTTCTCATCGGGA 798
2y 1066 AGACAGCATGCTCTTCAGGAGGAGCTCCAGGGCAAGAGGGTGTCTTGCTGTGCT 1125
3b 799 AGACAGCATGCTCTTCAGGAGGAGCTCCAGGGCAAGAGGGTGTCTTGCTGTGCT 858
2y 1126 TGAAGGGAACCTTGCATATCCCAAGTGCAGTCCCTCAGCTGTGGTGGCTTGCA 1185
3b 859 TGAAGGGAACCTTGCATATCCCAAGTGCAGTCCCTCAGCTGTGGTGGCTTGCA 918
2y 1186 TCCTGACTGATGTTCTCAGCCCTTGTCTGGCAAGAACCCAGAGCTCCCGAGTGTG 1245
3b 919 TCCTGACTGATGTTCTCAGCCCTTGTCTGGCAAGAACCCAGAGCTCCCGAGTGTG 978
2y 1246 ATACTAATAACCTCTTGGAGCAC 1269
3b 979 ATACTAATAACCTCTTGGAGCAC 1002

```

## RESULT 4

```

AK130234
LOCUS AK130234 1178 bp mRNA linear PRI 10-SEP-2003
DEFINITION Homo sapiens cDNA FLJ26724 fis, clone PNC05489.
ACCESSION AK130234
VERSION AK130234.1 GI:34526995
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Suzuki, O., Sasaki, N., Aotsuma, S., Shoji, T., Ichihara, T.,
Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,

```

Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiya, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kato, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsubo, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T., and Sugano, S.

NEDO human cDNA sequencing project

Unpublished  
2 (bases 1 to 1178)

Direct Submission  
Sugano, S. and Suzuki, Y.

Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail: flicdn@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

Location/Qualifiers  
1..1178  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PNC05489"  
/tissue\_type="pancreas"  
/clone\_lib="PNC"  
/note="cloning vector: pME18SFL3"

source

## ORIGIN

```

Query Match      72.4%; Score 927.6; DB 9; Length 1178;
Best Local Similarity 86.5%; Pred. No. 2.6e-219; Indels 154; Gaps 2;
Matches 1113; Conservative 0; Mismatches 19;

QY 30 GAGTTCGGGTGGGGAGCTTCACGAGGGAGGCTCTGTGAAAGAAACCCCTTTCTCT 89
DB 1 GAGTTCGGGTGGGGAGCTTCACGAGGGAGGCTCTGTGAAAGAAACCCCTTTCTCT 60
QY 90 CGCGGTGTCTACCCCTTTCTCCCATATCTGTTGGACATGAGCTGAGGACGGTGC 149
DB 61 CGCGGTGTCTACCCCTTTCTCCCATATCTGTTGGACATGAGCTGAGGACGGTGC 120
QY 150 GGGCGGTTCAGCCCTTTCGCGAGCTACGCGAGGAGGGCGGATGTTCTTGTTCGCG 209
DB 121 GGGCGGTTCAGCCCTTTCGCGAGCTACGCGAGGAGGGCGGATGTTCTTGTTCGCG 180
QY 210 TCCGCTTAGTGGCCGCTCCATTCGCGCGGTGTCCGATTTTAGGGGTAGGAGAGTG 269
DB 181 TCCGCTTAGTGGCCGCTCCATTCGCGCGGTGTCCGATTTTAGGGGTAGGAGAGTG 225
QY 270 TCAGCTTCAGGCAATCGCGAGGCGTGGCGGCCCATGGCGGCCCTGGAGGGCCCCGG 329
DB 226 ----- 225
QY 330 CTGGTACTGCTGTTTCAGCGGCAAGAGGAAATCCGGGAAGGACTTCGTGACCGAGCGCTG 389
DB 226 -----GGGAAGGACTTCGTGACCGAGCGCTG 252
QY 390 CAGAGCAGACTTGGAGCTGATGTCGTCTGTCTCGGCTCTCTGGTCCACTCAAGGAA 449
DB 253 CAGAGCAGACTTGGAGCTGATGTCGTCTGTCTCGGCTCTCTGGTCCACTCAAGGAA 312
QY 450 CAGTATGCTCAGGAGCATGGCTTGAATTCAGAGACTCTTGGACACAGGACCTTACAG 509
DB 313 CAGTATGCTCAGGAGCATGGCTTGAATTCAGAGACTCTTGGACACAGGACCTTACAG 372
QY 510 GAGGCTTTCCGAGGACATGATCGCTGGGAGGAGGAGAAACGCCAGGCTGACCCAGG 569
DB -----

```

Db 373 GAGGCTTTTGGAGGACATGATCCGCTGGGAGAGGAGAAAGCCAGGCTGACCCAGGC 432  
 Qy 570 TTCTTTTGGAGAGATTTGGAGGACATCTCCAGCCATCTGGCTGGCTGAGTGACACA 629  
 Db 433 TTCTTTTGGAGAGATTTGGAGGACATCTCCAGCCATCTGGCTGGCTGAGTGACACA 492  
 Qy 630 CGGAGAGTGTCTGACATCCAGTGGTTTCGGGAG----- 662  
 Db 493 CGGAGAGTGTCTGACATCCATGCCGCCCTGGAGGGCCCGCGGCTGGTACTGCTGTT 552  
 Qy 563 -----GCTATGGGCGCTGACGAGAGGCTCCGCTGGCTGGCTGAGGTT 703  
 Db 553 CAGCGGCAAGAGAAATCCGGAAGGATTCGTGACGAGACGCTCCGCTGGTGGGTT 612  
 Qy 704 GGAGCAGAGCCACAGCAGCGGCTGGGTTTTCAGCGCAGGAGGTGGACGATGCTGAGTC 763  
 Db 613 GGAGCAGAGCCACAGCAGCGGCTGGGTTTTCAGCGCAGGAGGTGGACGATGCTGAGTC 672  
 Qy 764 AGAATGTGCTGGACACATCTGGGAGATTTGACTGGGTCTACGAAACCATGAGTTGA 823  
 Db 673 AGAATGTGCTGGACACATCTGGGAGATTTGACTGGGTCTACGAAACCATGAGTTGA 732  
 Qy 824 ACAGCGCTGGAGGAGCAGTTGGAGAACTGATAGAAATTTATCCGCTCCAGACTTTAGTC 883  
 Db 733 ACAGCGCTGGAGGAGCAGTTGGAGAACTGATAGAAATTTATCCGCTCCAGACTTTAGTC 792  
 Qy 884 ACTAGGTTCTAGGAGTACTGGGCTGCTGAGTGGGGGTGGGCTGACTCTGCAGAAA 943  
 Db 793 ACTAGGTTCTAGGAGTACTGGGCTGCTGAGTGGGGGTGGGCTGACTCTGCAGAAA 852  
 Qy 944 TGGGGGTGTCCTCCGATCTGCGCCAGGTGAGAAACAGACAGGGGGGTCTAGATTTCTGA 1003  
 Db 853 TGGGGGTGTCCTCCGATCTGCGCCAGGTGAGAAACAGACAGGGGGGTCTAGATTTCTGA 912  
 Qy 1004 GGGGTTGGTGATATTTGGGCAAGCAGAAACCTCTGAGACCTCATTTCTCCATGGG 1063  
 Db 913 GGGGGTTGGTGATATTTGGGCAAGCAGAAACCTCTGAGACCTCATTTCTCCATGGG 972  
 Qy 1064 GAAGCAGCCATGCTCTTACAGAGAGACTCCAGAGCAAGAGGGGTCTTGGCTGTG 1123  
 Db 973 GAAGCAGCCATGCTCTTACAGAGAGACTCCAGAGCAAGAGGGGTCTTGGCTGTG 1032  
 Qy 1124 CTTGAAGGGGAAACCTGTCATATCCAGTGCAGTCCCTCAGCTGTGGTGGCTTG 1183  
 Db 1033 CTTGAAGGGGAAACCTGTCATATCCAGTGCAGTCCCTCAGCTGTGGTGGCTTG 1092  
 Qy 1184 CATCTGACTGATGTTCTCAGCCCTGTTTCTGGGCAAGACCCAGAGCTCCCGAGTGT 1243  
 Db 1093 CATCTGACTGATGTTCTCAGCCCTGTTTCTGGGCAAGACCCAGAGCTCCCGAGTGT 1152  
 Qy 1244 GGATACATAAATACCTCTTGGAGCAC 1269  
 Db 1153 GGATACATAAATACCTCTTGGAGCAC 1178

RESULT 5  
 LOCUS AY335584  
 DEFINITION Synthetic construct Homo sapiens phosphomevalonate kinase (PMVK)  
 ACCESSION AY335584  
 VERSION AY335584.1 GI:33303794  
 KEYWORDS FLI CDNA  
 SOURCE synthetic construct  
 ORGANISM artificial sequences  
 REFERENCE 1 (bases 1 to 579)  
 AUTHORS Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D., Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jenson,D., Harlow,E., LaBaer,J., and Brizuela,L.  
 TITLE Cloning of human full-length CDS FLEXGene kinases in recombinational vector system  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 579)  
 AUTHORS Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D., Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jenson,D., Harlow,E., LaBaer,J., and Brizuela,L.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUL-2003) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141-2023, USA  
 COMMENT This CDS clone is a part of a collection of human full-length expression clones generated by Harvard Institute of Proteomics. Each CDS has been cloned without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD in-fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-Dual vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.  
 FEATURES  
 source 1..579  
 /organism="synthetic construct"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:32630"  
 /clone="FLH000020.01L"  
 /clone\_lib="MGC clone templates"  
 /lab\_host="DH5alpha TI resistant"  
 /note="vector: pDNR-Dual"  
 gene 1..579  
 /gene="PMVK"  
 CDS 1..579  
 /gene="PMVK"  
 /note="Mutations: Stop->Leu"  
 /codon\_start=1  
 /transl\_table=11  
 /product="phosphomevalonate kinase"  
 /protein\_id="AAQ02411.1"  
 /db\_xref="GI:33303795"  
 /translation="MAPLGGAPRLVLLPSGKSKGKDFVTEALQSRGLADVCNVLRLS GLPKGYAGHGLNFQRLDSTYKAFKDKMIRWGEKQADPFCKIVGG:SQP IWLVSITRVSDIQWFRAYGAVTQTVRVVALEQSRQKRGWVTFPGVDDAESECLDN PGDFDNIENHGVQRLEQLENIETIRSL"  
 ORIGIN  
 Query Match 45.1%; Score 577.4; DB 12; Length 579;  
 Best Local Similarity 99.8%; Pred. No. 2.3e-132;  
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 303 ATGCCCCCGCTGGAGCGCCCCCGCGCTGCTGCTTTCAGCGGCAAGAGGAATCC 362  
 Db 1 ATGCCCCCGCTGGAGCGCCCCCGCGCTGCTGCTTTCAGCGGCAAGAGGAATCC 60  
 Qy 363 GGAAGAGCTTCTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATGTGTGCTGTC 422  
 Db 61 GGAAGAGCTTCTGACCGAGCGCTGCAGAGCAGACTTGGAGCTGATGTGTGCTGTC 120  
 Qy 423 CTCGGCTCTGCTCCACTCAAGGAACAGTGTCTCAGAGCATGGCTTGAATCTCCAG 482  
 Db 121 CTCGGCTCTGCTCCACTCAAGGAACAGTGTCTCAGAGCATGGCTTGAATCTCCAG 180  
 Qy 483 AGACTCTCTGACACACAGCAGCTTACAGAGGCGCTTTCGGAAGACATGATCCGCTGGGA 542  
 Db 181 AGACTCTCTGACACACAGCAGCTTACAGAGGCGCTTTCGGAAGACATGATCCGCTGGGA 240  
 Qy 543 GAGGAGAACCGCAGGCTGACCCAGGCTCTTTTTCAGGAAGATTCTGGAGGCGATCTCC 602  
 Db 241 GAGGAGAACCGCAGGCTGACCCAGGCTCTTTTTCAGGAAGATTCTGGAGGCGATCTCC 300  
 Qy 603 CAGCCATCTCTGGCTGAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGAG 662  
 Db 301 CAGCCATCTCTGGCTGAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGAG 360  
 Qy 663 GCCTATGGGCGCTGAGCGAGCGTCCGCTTGTAGCGTTGGAGCAGAGCCGACAGCAG 722  
 Db 361 GCCTATGGGCGCTGAGCGAGCGTCCGCTTGTAGCGTTGGAGCAGAGCCGACAGCAG 420

QY 723 CGGGCTGGGTGTTACGCCAGGGGTGGACGATGCTGAGTCAGAAATGTGGCCTGGACAAC 782  
 Db 421 CGGGCTGGGTGTTACGCCAGGGGTGGACGATGCTGAGTCAGAAATGTGGCCTGGACAAC 480  
 QY 783 TTCGGGACCTTGACTGGTTCATCGAGAACCATGAGTTGACAGCGCCCTGGAGGACAG 842  
 Db 481 TTCGGGACCTTGACTGGTTCATCGAGAACCATGAGTTGACAGCGCCCTGGAGGACAG 540  
 QY 843 TTGGAGAACCTTGATAGAAATTTATCGCTCCAGACTTTAG 881  
 Db 541 TTGGAGAACCTTGATAGAAATTTATCGCTCCAGACTTTAG 579

RESULT 6  
 AF026069  
 LOCUS 8862 bp DNA linear PRI 12-MAY-1999  
 DEFINITION Homo sapiens phosphomevalonate kinase (HUMPMKI) gene, partial cds.  
 ACCESSION AF026069  
 VERSION AF026069.1 GI:3647227

KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 8862)  
 AUTHORS Olivier, L.M., Chambliss, K.L., Gibson, K.M. and Krisans, S.K.  
 TITLE Characterization of phosphomevalonate kinase: chromosomal  
 localization, regulation, and subcellular targeting  
 J. Lipid Res. 40 (4), 672-679 (1999)

JOURNAL MEDLINE 99208737  
 PUBMED 10191291  
 REFERENCE 2 (bases 1 to 8862)  
 AUTHORS Chambliss, K.L. and Gibson, K.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-SEP-1997) Institute of Metabolic Diseases, Baylor  
 University Medical Center, 3812 Elm St., Dallas, TX 75226, USA

REFERENCE 3 (bases 1 to 8862)  
 AUTHORS Chambliss, K.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-SEP-1998) Pediatrics, University of Texas  
 Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX  
 75235-9063, USA

REMARK Sequence update by submitter  
 COMMENT On Sep 25, 1998 this sequence version replaced gi:3445541.

FEATURES  
 source

1. 8862  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="1p13-1q22-23"  
 <1. >7853

gene

mRNA join(<576..639,3813..3965,6507..6636,7717..>7853)

CDS /gene="HUMPMKI"

/evidence=experimental

/product="phosphomevalonate kinase"

/protein\_id="AAC60791.1"

/db\_xref="GI:3445542"

/translation="IGADVCAVLRSLGSLKEQYAEHGLNFORLLDSTVKAFKDM  
 IRWGEKQADPGFRCRIVSGISQPIWLVSDFTRVSDIQWFREAYGAVTQTVRVVAL  
 FQSRQQRGWFTPGVDAAEBCGLDNFGDFDVIENHGEVQRLEEQLENIIEFIRSL  
 "

ORIGIN

Query Match 41.2%; Score 528; DB 9; Length 8862;  
 Best Local Similarity 100.0%; Pred. No. 4e-120;  
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 743 AGGGGTGGACGATGCTGAGTCAGAAATGTGGCTGGACAACCTTCGGGACCTTTGACTGGGT 802  
 Db 7715 AGGGGTGGACGATGCTGAGTCAGAAATGTGGCTGGACAACCTTCGGGACCTTTGACTGGGT 7774  
 QY 803 CATCGAGAACCATGAGTTGAACAGCGCTCGAGGACGAGTTGGAGAACCTCATAGAAAT 862  
 Db 7775 CATCGAGAACCATGAGTTGAACAGCGCTCGAGGACGAGTTGGAGAACCTCATAGAAAT 7834  
 QY 863 TATCCGCTCCAGACTTTTACTCACTAGGTTCTAGGAGTGAAGTGGGCTGTGAGGTGG 922  
 Db 7835 TATCCGCTCCAGACTTTTACTCACTAGGTTCTAGGAGTGAAGTGGGCTGTGAGGTGG 7894  
 QY 923 GGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCGATCTCCGCGAGTGGAGGACAGA 982  
 Db 7895 GGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCGATCTCCGCGAGTGGAGGACAGA 7954  
 QY 983 CAGGGGGGCTCTAGATTCTGAGGGGGTGGTGGATATTGGGCAAGCGCAGGAAACCTCTGG 1042  
 Db 7955 CAGGGGGGCTCTAGATTCTGAGGGGGTGGTGGATATTGGGCAAGCGCAGGAAACCTCTGG 8014  
 QY 1043 AGACCTCATTTCTCCATGGGGAAGACAGCCATGCTCTTCAGAGGAGACTCAAGGGCA 1102  
 Db 8015 AGACCTCATTTCTCCATGGGGAAGACAGCCATGCTCTTCAGAGGAGACTCAAGGGCA 8074  
 QY 1103 MAGGAGGGTGTCTGGCTGTCTGAAGCGGAAACCTTCGCCATATCCCCAGTGCCAGTCC 1162  
 Db 8075 MAGGAGGGTGTCTGGCTGTCTGAAGCGGAAACCTTCGCCATATCCCCAGTGCCAGTCC 8134  
 QY 1163 CTTAGCCTGTGGTGGCTTGCATCTGAGTGGATTTCTAGCCCTTGTCTGGGCA 1222  
 Db 8135 CTTAGCCTGTGGTGGCTTGCATCTGAGTGGATTTCTAGCCCTTGTCTGGGCA 8194  
 QY 1223 GAACCCAGAGCTCCCGAGTGGATTAATAAACCCTCTTGAGGACA 1270  
 Db 8195 GAACCCAGAGCTCCCGAGTGGATTAATAAACCCTCTTGAGGACA 8242

RESULT 7

AL672265.1

WPCOMMENT

Sequence split into 6 fragments LOCUS AL672265 Accession AL672265

Fragment Name	Begin	End
AL672265.1	1	110000
AL672265.2	100001	210000
AL672265.3	200001	310000
AL672265.4	300001	410000
AL672265.5	400001	510000
AL672265.6	500001	534015

Continuation (2 of 6) of AL672265 from base 100001 (AL672265 Homo sapiens chromosome 1)

Query Match 41.2%; Score 528; DB 2; Length 110000;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-120;  
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 743 AGGGGTGGACGATGCTGAGTCAGAAATGTGGCTGGACAACCTTCGGGACCTTTGACTGGGT 802  
 Db 65400 AGGGGTGGACGATGCTGAGTCAGAAATGTGGCTGGACAACCTTCGGGACCTTTGACTGGGT 65459  
 QY 803 CATCGAGAACCATGAGTTGAACAGCGCTCGAGGACGAGTTGGAGAACCTCATAGAAAT 862  
 Db 65460 CATCGAGAACCATGAGTTGAACAGCGCTCGAGGACGAGTTGGAGAACCTCATAGAAAT 65519

QY 863 TATCCGCTCCAGACTTTTACTCACTAGGTTCTAGGAGTGAAGTGGGCTGTGAGGTGG 922  
 Db 65520 TATCCGCTCCAGACTTTTACTCACTAGGTTCTAGGAGTGAAGTGGGCTGTGAGGTGG 65579

QY 923 GGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCGATCTCCGCGAGTGGAGAACAGA 982  
 Db 65580 GGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCGATCTCCGCGAGTGGAGAACAGA 65639

QY 983 CAGGGGGGCTCTAGATTCTGAGGGGGTGGTGGATATTGGGCAAGCGCAGGAAACCTCTGG 1042

Db 65640 CAGGGGGGCTAGATTCTGAGGGGTTGGTGGATATGGGAGGAGGAACTCTG 65699  
 QY 1043 AGACCTCATTTTCTCATGAGGAGACGACGATCTTTCAGGAGGAGACTTCCAGGGCA 1102  
 Db 65700 AGACCTCATTTTCTCATGAGGAGACGACGATCTTTCAGGAGGAGACTTCCAGGGCA 65759  
 QY 1103 AAGGAGGGTGTCTTGGCTGTGCTTCAAGGCGGAAACCTCGCATATCCCGAGTGCAGTCC 1162  
 Db 65760 AAGGAGGGTGTCTTGGCTGTGCTTCAAGGCGGAAACCTCGCATATCCCGAGTGCAGTCC 65819  
 QY 1163 COTCAGCCTGTGGTGGCTTGCATCTGACGTGATGTTCTCAGCCCTGTGTTCTGGGCAA 1222  
 Db 65820 COTCAGCCTGTGGTGGCTTGCATCTGACGTGATGTTCTCAGCCCTGTGTTCTGGGCAA 65879  
 QY 1223 GAACCCAGAGCTCCCGAGTGTGATGTAATAAACCCTCTTGAGGACCA 1270  
 Db 65880 GAACCCAGAGCTCCCGAGTGTGATGTAATAAACCCTCTTGAGGACCA 65927

RESULT 8  
 AC021890 175826 bp DNA linear HTG 10-NOV-2000  
 LOCUS Homo sapiens chromosome 3 clone RP11-498A2, WORKING DRAFT SEQUENCE,  
 DEFINITION 22 unordered pieces.  
 AC021890  
 AC021890.12 GI:11128300  
 VERSION HTG, HTGS, PHASE1; HTGS\_DRAFT.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 175826)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oesman, F.R., Allen, C.,  
 Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Banks, J., Barbaria, J.,  
 Benton, J., Biemage, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
 Bowie, S., Brieva, M., Brown, M., Bryant, N.P., Buhay, C.,  
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
 Carter, N., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
 Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
 Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,  
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,  
 Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,  
 Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
 Louissegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M.,  
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S.,  
 Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Picken, R., Primus, E., Pu, L.,  
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,  
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoostari, N.,  
 Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Taylor, P., Telford, B., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Vera, V., Villalón, D., Vinson, R.,  
 Thomas, S., Usmani, K., Vazquez, L., Warren, R., Washington, C.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Wleciyk, K., Woodson, S.,  
 Watlington, S., Williams, G., Williamson, A., Wleciyk, K., Woodson, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.  
 and Gibbs, R.

TITLE  
 JOURNAL

2 (bases 1 to 175826)  
 Worley, K.C.  
 Direct Submission  
 Submitted (22-JAN-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 9, 2000 this sequence version replaced gi:9719697.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project information  
 Center project name: HMZE  
 Center clone name: RP11-498A2  
 ----- Summary Statistics  
 Sequencing vector: M13: L08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 144734 bases at least Q40  
 Consensus quality: 162873 bases at least Q30  
 Consensus quality: 169424 bases at least Q20  
 Estimated insert size: 170770; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 consists of 22 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1  
 21031: contig of 21031 bp in length  
 21032: gap of unknown length  
 21131: contig of 19344 bp in length  
 40476: gap of unknown length  
 40575: gap of unknown length  
 57617: contig of 17042 bp in length  
 57717: gap of unknown length  
 57718: contig of 17991 bp in length  
 75709: gap of unknown length  
 75809: contig of 13210 bp in length  
 89018: gap of unknown length  
 89019: contig of 11230 bp in length  
 100348: gap of unknown length  
 100449: contig of 10256 bp in length  
 110705: gap of unknown length  
 110804: contig of 12352 bp in length  
 123157: gap of unknown length  
 123257: contig of 9211 bp in length  
 132468: gap of unknown length  
 132568: contig of 8894 bp in length  
 141561: gap of unknown length  
 141562: contig of 5015 bp in length  
 146576: gap of unknown length  
 146677: contig of 4525 bp in length  
 151201: gap of unknown length  
 151202: contig of 4421 bp in length  
 151302: gap of unknown length  
 155723: contig of 4326 bp in length  
 155823: gap of unknown length  
 160149: contig of 2776 bp in length  
 160248: gap of unknown length  
 163024: contig of 2215 bp in length  
 163124: gap of unknown length  
 163135: contig of 2674 bp in length  
 165340: gap of unknown length  
 165439: contig of 1654 bp in length  
 168113: gap of unknown length  
 168214: contig of 1654 bp in length  
 168214: gap of unknown length  
 169867: gap of unknown length  
 169967: contig of 1613 bp in length  
 169968: contig of 1613 bp in length  
 169968: contig of 1613 bp in length



```

* 171581 171580: gap of unknown length
* 171681 172882: contig of 1202 bp in length
* 172983 172982: gap of unknown length
* 172983 174331: contig of 1449 bp in length
* 174332 174531: gap of unknown length
* 174532 175826: contig of 1295 bp in length.

FEATURES
    source
        1..175826
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="3"
            /clones="RP11-498A2"

ORIGIN
    Query Match 41.2%; Score 528; DB 2; Length 175826;
    Best Local Similarity 100.0%; Pred. No. 3.8e-120; Indels 0; Gaps 0;
    Matches 528; Conservative 0; Mismatches 0;

QY 743 AGGGTGGACGATCTGAGTCAGATGTGGCTGGCAACTTGGGACACTTGGGACCTTGAAGTGGT 802
DB 83982 AGGGTGGACGATCTGAGTCAGATGTGGCTGGCAACTTGGGACACTTGGGACCTTGAAGTGGT 84041
QY 803 CATCGAGAACCATGGTGTGAACAGCGCTGGAGGACGTTGGAGAACCTGATAGATT 862
DB 84042 CATCGAGAACCATGGTGTGAACAGCGCTGGAGGACGTTGGAGAACCTGATAGATT 84101
QY 863 TATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTGCTGAGGTGGG 922
DB 84102 TATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTGCTGAGGTGGG 84161
QY 923 GGTGGGGTGACTCTGCAAAATGGGGGTGTCCCGGATCTCCCGGAGTGTGGCCAGGAGTGAAGACAGA 982
DB 84162 GGTGGGGTGACTCTGCAAAATGGGGGTGTCCCGGATCTCCCGGAGTGTGGCCAGGAGTGAAGACAGA 84221
QY 983 CAGGGGGGTCTAGATCTGAGGGGGTGTGGTGATTTGGGCAAGGAGGAGTGAAGACAGA 1042
DB 84222 CAGGGGGGTCTAGATCTGAGGGGGTGTGGTGATTTGGGCAAGGAGGAGTGAAGACAGA 84281
QY 1043 AGACCTCATTTTCTCCATGGGGAAGACAGCATCTCTTACGAGGAGGAGATCCAAAGGCA 1102
DB 84282 AGACCTCATTTTCTCCATGGGGAAGACAGCATCTCTTACGAGGAGGAGATCCAAAGGCA 84341
QY 1103 AAGGAGGGTGTCTGGCTGTCTGAGGCGAAGACCTGCCATATCCCGAGTGCAGTCC 1162
DB 84342 AAGGAGGGTGTCTGGCTGTCTGAGGCGAAGACCTGCCATATCCCGAGTGCAGTCC 84401
QY 1163 CTTGAGCTGTGGTGGCTCTGATCCTGATGATGTTCTCAGGCCCTTGTCTTGGGCAA 1222
DB 84402 CTTGAGCTGTGGTGGCTCTGATCCTGATGATGTTCTCAGGCCCTTGTCTTGGGCAA 84461
QY 1223 GAACCCAGAGTCCCGAGTGGATCTAATAAACCTTGGAGCACA 1270
DB 84462 GAACCCAGAGTCCCGAGTGGATCTAATAAACCTTGGAGCACA 84509

RESULT 9
AL451085/C
LOCUS Human DNA sequence from clone RP11-307C12 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL451085 AC027440
VERSION AL451085.20 GI:19309454
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 182166)
AUTHORS Brown, A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

```

```

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 8, 2002 this sequence version replaced gi:19031396.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/Celegans/wormpep
This sequence was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP11-307C12 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-307C12. The true
right end of clone RP11-196D4 is at 89714 in this sequence.

FEATURES
    source
        1..182166
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-307C12"
            /clone_lib="RPCI-11.2"
            misc_feature
                61858..61932
                    /note="Sequence from overlapping clones RP11-196D4
                    (AL390204) and RP11-163D16 (AC034149). Assembly confirmed
                    by restriction digest."
            misc_feature
                72681..72974
                    /note="Sequence from overlapping clones RP11-196D4
                    (AL390204) and RP11-163D16 (AC034149). Assembly confirmed
                    by restriction digest."
            misc_feature
                81173..81260
                    /note="Sequence from overlapping clones RP11-196D4
                    (AL390204) and RP11-163D16 (AC034149). Assembly confirmed
                    by restriction digest."

ORIGIN
    Query Match 41.2%; Score 528; DB 9; Length 182166;
    Best Local Similarity 100.0%; Pred. No. 3.8e-120; Indels 0; Gaps 0;
    Matches 528; Conservative 0; Mismatches 0;

QY 743 AGGGTGGACGATCTGAGTCAGATGTGGCTGGCAACTTGGGACACTTGAAGTGGT 802
DB 53705 AGGGTGGACGATCTGAGTCAGATGTGGCTGGCAACTTGGGACACTTGAAGTGGT 53646
QY 803 CATCGAGAACCATGGTGTGAACAGCGCTGGAGGACGTTGGAGAACCTGATAGATT 862
DB 53645 CATCGAGAACCATGGTGTGAACAGCGCTGGAGGACGTTGGAGAACCTGATAGATT 53586
QY 863 TATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTGCTGAGGTGGG 922
DB 53585 TATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTGCTGAGGTGGG 53526
QY 923 GGTGGGGTGACTCTGCAAAATGGGGGTGTCCCGGATCTCCCGGAGTGTGAAGACAGA 982
DB 53525 GGTGGGGTGACTCTGCAAAATGGGGGTGTCCCGGATCTCCCGGAGTGTGAAGACAGA 53466
QY 983 CAGGGGGGTCTAGATCTGAGGGGGTGTGGTGATTTGGGCAAGGAGGAGTGAAGACAGA 1042

```





```
Db 134171 GGTGGGCTGACTCTGCAAAATGGGGGTGTCCCGATCCTGGCCGAGGTGAGGAACAGA 134112
2y 983 CAGGGGGTCTAGATTCTGAGGGGGTGGTGGATATTGGGCAAGCAGACCACTCTGG 1042
Db 134111 CAGGGGGGTCTAGATTCTGAGGGGGTGGTGGATATTGGGCAAGCAGACCACTCTGG 134052
2y 1043 AGACCTCATTTTCTCCATGGGGAAGACAGACCATGCTCTTCAGGAGGAGATCCCAAGGGCA 1102
Db 134051 AGACCTCATTTTCTCCATGGGGAAGACAGACCATGCTCTTCAGGAGGAGATCCCAAGGGCA 133992
2y 1103 AAGAGAGGTGCTTGGTGTGCTTGAAGGCGAAGACCTGCGCATATCCCAAGTGGCAGTCC 1162
Db 133991 AAGAGAGGTGCTTGGTGTGCTTGAAGGCGAAGACCTGCGCATATCCCAAGTGGCAGTCC 133922
2y 1163 CCTCAGCCTGTGTGGCTTGTGATCTTCAGTGGATGTCTCAGCCCTTGTCTGGGCA 1222
Db 133931 CCTCAGCCTGTGTGGCTTGTGATCTTCAGTGGATGTCTCAGCCCTTGTCTGGGCA 133872
2y 1223 GAACCCAGAGCTCCCGAGTGGATCTTAATAACCTTTGGAGCACA 1270
Db 133871 GAACCCAGAGCTCCCGAGTGGATCTTAATAACCTTTGGAGCACA 133824

RESULT 11
LOCUS AX886270 465 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2133 from Patent EP1033401.
ACCESSION AX886270
VERSION AX886270.1 GI:40043263
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Dumas Milne Edwards, J.B., Duclat, A. and Giordano, J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 2133 06-SEP-2000;
GENSET (PR)
FEATURES
source Location/Qualifiers
1..465
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
57..>464
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAP00408.1"
/db_xref="GI:40043264"
/translation="MAPLGGAPRLVLLFSXKRKSGKDFVTZALQSLRGADVCAVRLS
GPLKEQVAQEHQNFLLDTSTYKEAFKRMIRWBEKQAXXFPCKRIVEGISQ
IWLVSDFTRVSDIQWFREAVGAVTQTVRVVVALER"
CDS
ORIGIN
Query Match 35.8%; Score 459.2; DB 6; Length 465;
Best Local Similarity 98.5%; Pred. No. 5.1e-103;
Matches 458; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 247 GATTTTAGGGGTAGGAGAGTGTCTCAGGATCGCGAGCGGTGGCGGCCCATGG 306
Db 1 GATTTTAGGGGTAGGAGAGTGTCTCAGGATCGCGAGCGGTGGCGGCCCATGG 60
QY 307 CCCCGCTGGAGGCGCCCGCGGTGTGCTGTTTCAGCGGCAAGAGAAATCCGGGA 366
Db 61 CCCCGCTGGAGGCGCCCGCGGTGTGCTGTTTCAGCGGCAAGAGAAATCCGGGA 120
QY 367 AGGACTTCGTGACCGAGCGGTGCGAGCAGACTTGGAGCTGATGTCTGTCTCTCC 426
Db 121 AGGACTTCGTGACCGAGCGGTGCGAGCAGACTTGGAGCTGATGTCTGTCTCTCC 180
QY 427 GGCTCTCTGGTCCACTCAAGGAACAGTATGTTCAGGAGCATGGCTTGAACCTCCAGAGC 486
Db 181 GGCTCTCTGGTCCACTCAAGGAACAGTATGTTCAGGAGCATGGCTTGAACCTCCAGAGC 240
```

```
QY 487 TCCTGGACACGACACCTACAGAGGCGCTTTTCGGAAGGACATGCTCGCTGGGAGAGG 546
Db 241 TCCTGGACACGACACCTACAGAGGCGCTTTTCGGAAGGACATGCTCGCTGGGAGAGG 300
QY 547 AGAAACGCGAGGCTGACCCAGGCTTCTTTTCGAGGAAGATTGTGGAGGCGATCTCCAGC 606
Db 301 AGAAACGCGAGGCTGACCCAGGCTTCTTTTCGAGGAAGATTGTGGAGGCGATCTCCAGC 360
QY 607 CCATCTGGCTGTGTAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGGAGGCT 666
Db 361 CCATCTGGCTGTGTAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGGAGGCT 420
QY 667 ATGGGGCGGTGACGAGCAGCGTCCCGCTTGTAGCGTTGGAGCAGA 711
Db 421 ATGGGGCGGTGACGAGCAGCGTCCCGCTTGTAGCGTTGGAGCAGA 465

RESULT 12
LOCUS BD025880 465 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD025880
VERSION BD025880.1 GI:22567103
KEYWORDS JP 2001269182-A/2126.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Edwards, J.B.D.M., Duclat, E. and Jordan, J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 2126 02-OCT-2001;
GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/2126
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAT, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..465
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 35.8%; Score 459.2; DB 6; Length 465;
Best Local Similarity 98.5%; Pred. No. 5.1e-103;
Matches 458; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 247 GATTTTAGGGGTAGGAGAGTGTCTCAGGATCGCGAGCGGTGGCGGCCCATGG 306
Db 1 GATTTTAGGGGTAGGAGAGTGTCTCAGGATCGCGAGCGGTGGCGGCCCATGG 60
QY 307 CCCCGCTGGAGGCGCCCGCGGTGTGCTGTTTCAGCGGCAAGAGAAATCCGGGA 366
Db 61 CCCCGCTGGAGGCGCCCGCGGTGTGCTGTTTCAGCGGCAAGAGAAATCCGGGA 120
QY 367 AGGACTTCGTGACCGAGCGGTGCGAGCAGACTTGGAGCTGATGTCTGTCTCTCC 426
Db 121 AGGACTTCGTGACCGAGCGGTGCGAGCAGACTTGGAGCTGATGTCTGTCTCTCC 180
QY 427 GGCTCTCTGGTCCACTCAAGGAACAGTATGTTCAGGAGCATGGCTTGAACCTCCAGAGC 486
```

Db 181 GGCTCTCTGCTCCACTCAAGGACACAGTATGCTCAGGAGCATGGCTTGAATCTCCAGAGAC 240

Qy 487 TCTTGGACACACAGACCTTCAAGAGAGGCTTTCGAAGAGCATGATCGCTGGGAGAGG 546

Db 241 TCTTGGACACACAGACCTTCAAGAGAGGCTTTCGAAGAGCATGATCGCTGGGAGAGG 300

Qy 547 AGAAAGCCGAGGCTGACCCAGGCTTCTTTGAGGAGAGATGTTGGAGGAGCATCTCCAGC 606

Db 301 AGAAAGCCGAGGCTGNNCAGGCTTCTTTGAGGAGAGATGTTGGAGGAGCATCTCCAGC 360

Qy 607 CCATCTGGCTGTGAGTGACACACGAGAGTGCTTCGACATCCAGTGGTTTGGAGAGGCTT 666

Db 361 CCATCTGGCTGTGAGTGACACACGAGAGTGCTTCGACATCCAGTGGTTTGGAGAGGCTT 420

Qy 667 ATGGGCGCTGAGCAGAGCGTCCGCTTGTAGCGTTGGAGCAGA 711

Db 421 ATGGGCGCTGAGCAGAGCGTCCGCTTGTAGCGTTGGAGCGGA 465

RESULT 13

BC028659

LOCUS

DEFINITION

Mus musculus RIKEN cDNA 1110011E12 gene, mRNA (cdna clone OCT-2003 IMAGE:1247653), complete cds.

ACCESSION

BC028659

VERSION

BC028659.1

KEYWORDS

GI:20306533

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 1008)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

REFERENCE

2 (bases 1 to 1008)

Strausberg, R.

Direct Submission

Submitted (23-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NTH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)

Tissue Procurement: Marcello Bento Soares, Ph.D.

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettner, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 66 Row: h Column: 3.

Location/Qualifiers

1. .1008

# FEATURES source

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="MGC:41584 IMAGE:1247653"

/issue\_type="Mammary gland, lactating mouse"

/clone\_lib="Soares\_mammary\_gland\_NMLMG"

/lab\_host="DH10B"

/note="Vector: pT73-Pac"

## gene

1. .1008

/gene="1110011E12Rik"

/note="synonym: 2900002L22R1k"

/db\_xref="LocusID:68603"

/db\_xref="MGI:1915853"

## CDS

266..619

/codon\_start=1

/product="1110011E12Rik protein"

/protein\_id="AAH28659.1"

/db\_xref="GI:20306534"

/db\_xref="LocusID:68603"

/translation="MICNGEQKQADPGFCKIVEGVSOPIWLYSDTRTSDIOWFQ EAYGAVIQIRVVASEQSRQQRGWFTPOVDDAEESECLDNFNGFNDWVIEHGDDEQL EDQLEHLGLIOAKL"

## misc\_feature

265..427

/note="P-nevallo kinase; Region: Phosphomevalonate kinase. Phosphomevalonate kinase (EC:2.7.4.2) catalyzes the phosphorylation of 5-phosphomevalonate into 5-diphosphomevalonate, an essential step in isoprenoid biosynthesis via the mevalonate pathway. This family represents the animal type of the enzyme. The other is the ER8 type, found in plants and fungi, and some bacteria (see pfam00288)"

/db\_xref="CDD:pfam04275"

## ORIGIN

Query Match 28.1%; Score 359.6; DB 10; Length 1008;

Best Local Similarity 72.4%; Pred. No. 2.7e-78;

Matches 552; Conservative 0; Mismatches 194; Indels 16; Gaps 6;

Qy 456 GCTCAGGAGCATGGCTTGAACCTTCCAGAGACTCTCTGGACACACACACAGGAGGCC 515

Db 194 GCAAGAGGATGGCTTGGACTTCCAGAGACTTCTGGATCGAGACCTACAGGAGACC 253

Qy 516 TTTGGAGGACATGATTCCTGGGAGAGGAGAAACCGCAGCTGACCCAGGCTTCTTT 575

Db 254 TATCGGAGGACATGATCTCTGGGGGAGCAGAGCGCCAGCAGCCAGGCTTCTTC 313

Qy 576 TGCAGGAGATTTGGAGGGCATCTCCAGCCCATCTGGCTGTGAGTGACACACGAGA 635

Db 314 TCCGGAGAGATTGGAGGCGTGTCCAGCTATCTGGCTGTGAGTGACACACGAGG 373

Qy 636 GTGCTCAGATCCAGTGGTTTCGGAGGCGCTATGGGGCGGTGACGAGAGCGTCCGGTT 695

Db 374 ACATCTCAGATCCAGTGGTTTCAGAGGCGCTATGGGGCTGTGATACAGACAGTCCGAGTA 433

Qy 696 GTACGCTTGGAGCAGAGCCGACACAGCGGGGCTGGGTGTTCACGCCAGGGGTGGACGAT 755

Db 434 GTGCGCTCGGAGCAGAGTCCAGACACACGCGGGGTGGGTGTTCACACCGGGGTGGACGAT 493

Qy 756 GCTGAGTCAGAAATGTGGCTGGACAACTTCGGGAGCTTTGACCTGGGTTCATCGAGACCAT 815

Db 494 GCTGAGTCAGAGTGTGGTCTGGACAACTTTGGGAACTTTGACCTGGGTTCATTTGAGAACCCAC 553

Qy 816 GGACTTCAACAGCGCTGGAGGAGCGATTCGAGAACCTGATAGAAATTTATCCGCTCCAGA 875

Db 554 GGAGATGAGCAGTCCCTGGAGGATCAGCTGGAGACCTGCTGGGATTTATTCAGGCCAAA 613

Qy 876 CTTTAGTCACTAGGTTCTAGGAGTGAGCTGGGGCTGCTGAGGTGGGGGTGGGGCTGACT 935

Db 614 CTTTAGTGATGAGGCTTGAGGGACGAGAGAGACTGATGGCTGCGCAACAG-TGGTT 672  
 Qy 936 CTGCAAAATGGGGTGTC--CCGATCTGGCCGAGGTGAGNACAGACAGGGGGGTC 993  
 Db 673 CTGCAGATGGGTGCCAGTCCAGCCGAGGTTCAGAAACAGACAGACAGTCTGGCTT 732  
 Qy 994 TAGA-----TTCAGGGGGTGGTGATATGGGCAAGGCAAGCAACCTCTGGAGACC 1047  
 Db 733 GCCAGAGTCTGGGCAGGATCTCGTGAGTAGTGGCAATATAAGAACTCTGGTG-- 789  
 Qy 1048 TCATTTCTCCATGGGAGACAGCCATGCTCTCAGAGAGAGACTCCAAGGGCAAGGA 1107  
 Db 790 TTGTGTTTCCCTGGAGAGGACATCTGCTCTTTTAGGGC---ACTTATAGGGCAAGGC 846  
 Qy 1108 GGGTGTCTTGCTGTGCTGTGAAGCGCAAAACCTGCCATATCCCGAGTCCCGCTCA 1167  
 Db 847 AGGCGCATGATTTTGCTGAGGAGGAGACATTCGCACACCATGATGGCAGGCTGCAG 906  
 Qy 1168 GCCTGTGGTGGCCCTGGCAT-CTGACTGGATGTTCTCAGCCC 1208  
 Db 907 AGCTGTGATGTGTGTGTTTCCCTGTGTACGCTCTCAGCCC 948

RESULT 14  
 G22398/c 422 bp DNA linear STS 31-MAY-1996  
 LOCUS human STS WI-14846, sequence tagged site.  
 DEFINITION G22398  
 ACCESSION G22398  
 VERSION G22398.1 GI:1342724  
 STS; STS sequence; primer; sequence tagged site.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 422)  
 AUTHORS Hudson, T.  
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped STS  
 JOURNAL Unpublished (1995)  
 COMMENT

Contact: Thomas Hudson  
 Whitehead Institute/MIT Center for Genome Research  
 Whitehead Institute for Biomedical Research  
 9 Cambridge Center, Cambridge MA 02142 USA  
 Tel: 617 252 1900  
 Fax: 617 252 1902  
 Email: thudson@genome.wi.mit.edu

Primer A: AGGTTTATTAGTATCCACACTGGG  
 Primer B: TGCTTGGCTGTGCTTGAAG  
 STS size: 150  
 PCR Profile:

Presoak:  
 Denaturation:  
 Annealing: 56 degrees C  
 Polymerization:  
 PCR Cycles: 35  
 Thermal Cycler:

Protocol:  
 Template: 10 ng  
 Primer: each 5 pM  
 dNTPs: each 4 mM  
 Tag Polymerase: 0.025 units/ul  
 Total Vol: 20 ul

Buffer:  
 MgCl2: 1.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 9.3

Derived from dbEST (genbank accession H09819).

FEATURES  
 source  
 1..422  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="730.7 cR from top of Chr1 linkage group"  
 STS  
 primer\_bind  
 23..46  
 primer\_bind  
 complement (153..172)  
 ORIGIN

Query Match 26.0%; Score 333.6; DB 11; Length 422;  
 Best Local Similarity 97.0%; Pred. No. 7.7e-72;  
 Matches 360; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
 Qy 913 CTGAGGTGGGGTGGGGTGAATCTGCAAAATGGGGGTGTCCCCCGATCTCTGGCCGAGGT 972  
 Db 371 CTNAGGTGGGGTGGGGTGAATCTGCAAAATGGGGGTGTCCCCCGATCTCTGGCCGAGGT 312  
 Qy 973 GAGGAACAGACAGGGGGGT-CTAGATTCTGAGGGGGTGTGGGATATGGCAGAGGAG 1031  
 Db 311 GAGGAACAGACAGGGGGGTNCTAGATTCTGAGGGGGTGTGGGATATTTGGCAAGGCAG 252  
 Qy 1032 GAAACCTCTGGAGACCTCATTTTCTCCATGGGAAAGACAGCCATCTCTTCAGGAGGAGA 1091  
 Db 251 GAAACCTCTGGAGACCTCATTTTCTCCATGGGAAAGACAGCCATCTCTTCAGGAGGAGA 192  
 Qy 1092 CTCGAAGGCAAGGAGGGGTGTCTGGCTGTGCTTGAAGGGGAAACCTGCCATATCCCC 1151  
 Db 191 CTCGAAGGCAAGGAGGGGTGTCTGGCTGTGCTTGAAGGGGAAACCTGCCATATCCCC 132  
 Qy 1152 AGTGCCAGTCCCTCAGCCTGTGGTGGCTTGCATCTCCTGACATGGATGTTCTCAGCCCTT 1211  
 Db 131 AGTGCCAGTCCCTCAGCCTGTGGTGGCTTGCATCTCCTGACATGGATGTTCTCAGCCCTT 72  
 Qy 1212 GTT-CTGGCAAGAACCCAGAGCTCCCGAGTGTGATATAATAAACCTTTGGAGCACA 1270  
 Db 71 GTTCTGGCAAGAACCCAGAGCTCCCGAGTGTGATATAATAAACCTTTGGAGCACA 12  
 Qy 1271 AAAAAAAAAA 1281  
 Db 11 AAAAAAAAAA 1

RESULT 15  
 AR415566 419 bp DNA linear PAT 18-DEC-2003  
 LOCUS Sequence 3203 from patent US 6639063.  
 DEFINITION AR415566  
 ACCESSION AR415566  
 VERSION AR415566.1 GI:40170676  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclasseified.

REFERENCE 1 (bases 1 to 419)  
 AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.  
 TITLE EST's and encoded human proteins  
 JOURNAL Patent: US 6639063-A 3203 28-OCT-2003;  
 FEATURES  
 Location/Qualifiers  
 1..419  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 24.8%; Score 317.2; DB 6; Length 419;  
 Best Local Similarity 85.6%; Pred. No. 9e-68;  
 Matches 416; Conservative 0; Mismatches 3; Indels 67; Gaps 3;  
 Qy 229 CATTCGCGCGGTGTCCTGATTTTAGGGGTAGGAGAGTGTACGCTTCAGGCATCGCA 288  
 Db 1 CATTCGCGCGGTGTCCTGATTTTAGGGGTAGGAGAGTGTACGCTTCAGGCATCGCA 60  
 Qy 289 GCGGTGCGCGCCCATGCGCCCGCGGTGGAGGCGCCCGCGGTGTGTACTGTCTTACGG 348

```
Db      |||||
61  GGCGTGGCGGCCCATGGGCC--GCTGGAGGGGCCCGCGGTGTTACTGCTGTTCAGCG 118
QY      |||||
349  GCAAGAGGAAATCCGGGAAGGACTTCGTGACCGAGCGCTGCAGAGCAGACTTGGAGCTG 408
Db      |||||
119  GCAAGAGGAAATCCGGGAAGGACTTCGTGACCGAGCGCTGCAGAG----- 164
QY      |||||
409  ATGTCTGTCTGTCTCTCGGCTCTCTGTGTCCTCAAGGAACAGTATGCTCAGGAGCATG 468
Db      |||||
165  -----CAGGAGCATG 174
QY      |||||
469  GCTTGAACCTCCAGAGACTCTCTGGACACCCAGCACCTACAAAGGAGGCTTTCGGAAGGACA 528
Db      |||||
175  GCTTGAACCTCCAGAGACTCTCTGGACACCCAGCACCTACAAAGGAGGCTTTCGGAAGGACA 234
QY      |||||
529  TGATCCGCTGGGAGAGGAGAAACGCCAGGCTGACCCAGGCTTCTTTTCAGGAAAGATTG 588
Db      |||||
235  TGATCCGCTGGGAGAGGAGAAACGCCAGGCTGACCCAGGCTTCTTTTCAGGAAAGATTG 294
QY      |||||
589  TGGAGGGCATCTCCAGAGCCATCTGGCTGGTGAGTGACACAGGAGAGTGTCTGACATCC 648
Db      |||||
295  TGGAGGGCATCTCCAGAGCCATCTGGCTGGTGAGTGACACAGGAGAGTGTCTGACATCC 354
QY      |||||
649  AGTGGTTTCGGAGGCGCTATGGGGCCGTGACAGCAGCGGTCCGCGTTGTAGCGTTGGAGC 708
Db      |||||
355  AGTGGTTTCGGAGGCGCTATGGGG--CGTGACGAGACGCGTCCGCGTTGTAGCGTTGGAGC 413
QY      |||||
709  AGAGCC 714
Db      |||||
414  AGAGCC 419
```

Search completed: March 11, 2004, 03:35:24  
Job time : 3271.98 secs

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a







4141	ATCCATTGATTTAA	CCAACTTCTCGTGGT	TACTCACC	TGTGAAAAAT	GAGAGATGT	4200	
4201	TGCAGAGTTGT	CAAAATTTCTCTT	CAGACAAT	CAGTGGTCTT	TAAACAAATCTCTCGCTGG	4260	
4201	TGCAGAGTTGT	CAAAATTTCTCTT	CAGACAAT	CAGTGGTCTT	TAAACAAATCTCTCGCTGG	4260	
4261	TACAGAAATATG	TAGTCAGTGTCT	CAGTGTCT	CAGAAACAACAT	GAGAGCACACCTCTTAG	4320	
4261	TACAGAAATATG	TAGTCAGTGTCT	CAGTGTCT	TACGAAACAACAT	GAGAGCACACCTCTTAG	4320	
4321	AGGAAGACGAAA	CAAGGCTTGATTT	CCCAACTGGCA	TTGACTTTTCT	TGATATTA	4380	
4321	AGGAAGACGAAA	CAAGGCTTGATTT	CCCAACTGGCA	TTGACTTTTCT	TGATATTA	4380	
4381	CAACTCTTTTAC	TGTGCATGTGAT	TGCTCTCGAG	CCACCATCAC	TGGGTACAGGATCGG	4440	
4381	CAACTCTTTTAC	TGTGCATGTGAT	TGCTCTCGAG	CCACCATCAC	TGGGTACAGGATCGG	4440	
4441	CCATCATCCGAG	CACTTCAGTGGG	AGATCGT	CAGAAGATCGG	GGTCCCACTCTCGGAA	4500	
4501	TTCCATCACCT	CACCACTCACT	CCAGCA	CAGAGTATGGT	CAGCATGGTGTCT	4560	
4561	TAAATGGCAGAG	AGGAAGTCCCTTA	TGATTTGCC	CAACATCA	AGTTTCTGATGTTCC	4620	
4561	TAAATGGCAGAG	AGGAAGTCCCTTA	TGATTTGCC	CAACATCA	AGTTTCTGATGTTCC	4620	
4621	GAGGACCTTGG	AAAGTTGTGCTCG	CAACCCCA	CCAGCTACTGAT	CAGCTGGGATGCTCC	4680	
4621	GAGGACCTTGG	AAAGTTGTGCTCG	CAACCCCA	CCAGCTACTGAT	CAGCTGGGATGCTCC	4680	
4681	TGCTGTCA	GATGAGATATTA	CAGGATCACTT	TACGAGAAAC	GAGGAGGAATATG	4740	
4681	TGCTGTCA	GATGAGATATTA	CAGGATCACTT	TACGAGAAAC	GAGGAGGAATATG	4740	
4741	CCAGGAGTTCA	CTGTGCTGGG	AGCAAGTCTA	CAGGTTACCAT	CAGCGGGCTTAAAC	4800	
4741	CCAGGAGTTCA	CTGTGCTGGG	AGCAAGTCTA	CAGGTTACCAT	CAGCGGGCTTAAAC	4800	
4801	AGTTGATATTA	CCATCACTGTGT	ATGCTGTCACT	TGGCTGGG	AGACAGCCCCCAGCAG	4860	
4801	AGTTGATATTA	CCATCACTGTGT	ATGCTGTCACT	TGGCTGGG	AGACAGCCCCCAGCAG	4860	
4861	CAAGCCAAATTT	CCATTAATTA	TCGCAACAG	AAATTTGACAA	ACCATCCAGATGCAAGTGAC	4920	
4861	CAAGCCAAATTT	CCATTAATTA	TCGCAACAG	AAATTTGACAA	ACCATCCAGATGCAAGTGAC	4920	
4921	CGATGTT	CAGGACAA	CAGCATATAGT	GTCAAGTGGCT	TGCCTTCAAGTTCCCTGTTACTGG	4980	
4921	CGATGTT	CAGGACAA	CAGCATATAGT	GTCAAGTGGCT	TGCCTTCAAGTTCCCTGTTACTGG	4980	
4981	TTACAGAGTAA	CCACCACTCCCA	AAAAATGGAC	CCAGGACCAACAA	AAACTAAAACTGCAGG	5040	
4981	TTACAGAGTAA	CCACCACTCCCA	AAAAATGGAC	CCAGGACCAACAA	AAACTAAAACTGCAGG	5040	
5041	TCCAGATCA	AAACAGAAATG	ACTATTG	NAGGCTTTCAG	CCCCACAGTGGGATGCTGGTTAG	5100	
5041	TCCAGATCA	AAACAGAAATG	ACTATTG	NAGGCTTTCAG	CCCCACAGTGGGATGCTGGTTAG	5100	
5101	TGTCATATGCT	CAGAAATCCA	AGCGGAGAGT	CAGCTCTCG	TCTGGTTTCAGACTGCAGTAACCCAA	5160	
5101	TGTCATATGCT	CAGAAATCCA	AGCGGAGAGT	CAGCTCTCG	TCTGGTTTCAGACTGCAGTAACCCAA	5160	
5161	CATTGATCG	CCCTTAAAGG	ACTTGC	ATTA	CTGATGTGGATGTCGATTC	CAATCAAAATTCG	5220
5161	CATTGATCG	CCCTTAAAGG	ACTTGC	ATTA	CTGATGTGGATGTCGATTC	CAATCAAAATTCG	5220
5221	TTGGAAGCC	CCACAGG	GGCAAGTTTCC	AGGTAC	AGGCTGACTTCTCGAGCCCTGAGGA	5280	
5221	TTGGAAGCC	CCACAGG	GGCAAGTTTCC	AGGTAC	AGGCTGACTTCTCGAGCCCTGAGGA	5280	

5281	TGGAAATCAATGAGCTATTCCCTGCACCTGATGGTGAAAGAACACTGCAGAGCTGCAGG	5340
5281	TGGAAATCAATGAGCTATTCCCTGCACCTGATGGTGAAAGAACACTGCAGAGCTGCAGG	5340
5341	CCTCAGACCGGGTTCAGTAGTACACAGTCAAGTGTGGTTGCTTGCACGATGATATGGAGAG	5400
5341	CCTCAGACCGGGTTCAGTAGTACACAGTCAAGTGTGGTTGCTTGCACGATGATATGGAGAG	5400
5401	CCAGCCCTGANTTGGAAACCCAGTTCACAGCTATTCCCTGCACCACTGCAGCTGAAGTTCAC	5460
5401	CCAGCCCTGANTTGGAAACCCAGTTCACAGCTATTCCCTGCACCACTGCAGCTGAAGTTCAC	5460
5461	TCAGGTCAACCCCAACAGCCTGAGCGCCCAAGTGGACACCAACCTGTTGAGTCTCACTGG	5520
5461	TCAGGTCAACCCCAACAGCCTGAGCGCCCAAGTGGACACCAACCTGTTGAGTCTCACTGG	5520
5521	ATATCGAGTGGGGTGACCCCAAGAGAAAGACCGGACCAATGAAGAAATCAACCTTGC	5580
5521	ATATCGAGTGGGGTGACCCCAAGAGAAAGACCGGACCAATGAAGAAATCAACCTTGC	5580
5581	TCCTGACAGCTCATCCGTGGTTGATCAGGACTTATGTGTGGCCACCAAAATATGAAGTGAG	5640
5581	TCCTGACAGCTCATCCGTGGTTGATCAGGACTTATGTGTGGCCACCAAAATATGAAGTGAG	5640
5641	TGCTATGCTTTAAGACACTTTTGACAAGCAGACACAGCTCAGGGAGTGTGCACCACTCT	5700
5641	TGCTATGCTTTAAGACACTTTTGACAAGCAGACACAGCTCAGGGAGTGTGCACCACTCT	5700
5701	GGAGAAATGTCAGCCCCCAAGAAAGGGCTCGTGTGACAGATGCTACTGAGACCCACATCAC	5760
5701	GGAGAAATGTCAGCCCCCAAGAAAGGGCTCGTGTGACAGATGCTACTGAGACCCACATCAC	5760
5761	CAATAGCTGGAGAACCAAGACTCGACAGCATCACTGGCTCCAAAGTTGATGCCGTTCCAGC	5820
5761	CAATAGCTGGAGAACCAAGACTCGACAGCATCACTGGCTCCAAAGTTGATGCCGTTCCAGC	5820
5821	CAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAAGCTACACCATCAC	5880
5821	CAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAAGCTACACCATCAC	5880
5881	TGGCTTACAAACAGGCACTGACTACAGATCTACTGCTGATACCTTGAATGATCAATGCTCG	5940
5881	TGGCTTACAAACAGGCACTGACTACAGATCTACTGCTGATACCTTGAATGATCAATGCTCG	5940
5941	GAGCTCCCTGTGGTTCATCGACGCCTCCACTGCAATGTATGACCAATCAACCTGCGTTT	6000
5941	GAGCTCCCTGTGGTTCATCGACGCCTCCACTGCAATGTATGACCAATCAACCTGCGTTT	6000
6001	CTGGCCACACACCCAAATTCCTTGCTGGTATCATGCGACCGCCACCTGCCAGGATTAC	6060
6001	CTGGCCACACACCCAAATTCCTTGCTGGTATCATGCGACCGCCACCTGCCAGGATTAC	6060
6061	CGGCTACATCATCAAGTATGAGAAGCCTGGGTCTCTCCAGAGAAGTGGTCTCCCTCGGCC	6120
6061	CGGCTACATCATCAAGTATGAGAAGCCTGGGTCTCTCCAGAGAAGTGGTCTCCCTCGGCC	6120
6121	CGCCCTGGTGTACAGAGGCTACTATTACTGGCTGTGAAACCGGGAACCGAATATACAAT	6180
6121	CGCCCTGGTGTACAGAGGCTACTATTACTGGCTGTGAAACCGGGAACCGAATATACAAT	6180
6181	TTATGTCTATTGCCCTGAAGAAATTAATCAGAAGAGCGAGCCCTGATTGGAAGGAAAGAGAC	6240
6181	TTATGTCTATTGCCCTGAAGAAATTAATCAGAAGAGCGAGCCCTGATTGGAAGGAAAGAGAC	6240
6241	AGACGAGCTTCCCAACTGTGTACCTTCCACACCCCAATCTTCATGGACCCAGAGATCTT	6300
6241	AGACGAGCTTCCCAACTGTGTACCTTCCACACCCCAATCTTCATGGACCCAGAGATCTT	6300
6301	GGATGTTCCCTTCCACAGTTCAAAGACCCCTTTCGTCACCCACCTCGGATATGACTGG	6360
6301	GGATGTTCCCTTCCACAGTTCAAAGACCCCTTTCGTCACCCACCTCGGATATGACTGG	6360

6361 AAATGGTATTACAGTCTCTGGGCACTTCTGGTCAGCAACCCAGTGTGGGCAACAATGAT 6420  
 6361 AAATGGTATTACAGTCTCTGGGCACTTCTGGTCAGCAACCCAGTGTGGGCAACAATGAT 6420  
 6421 CTTTGGAGGAACATGGTTTGGAGGAGCAACACCGCCCAACAACCGGCAACCCCAATAGGCA 6480  
 6421 CTTTGGAGGAACATGGTTTGGAGGAGCAACACCGCCCAACAACCGGCAACCCCAATAGGCA 6480  
 6481 TAGCCCAAGACCAATACCCGCGAATCTAGGCAAGAAGCTCTCTCTCAGACAACCAATCTC 6540  
 6481 TAGCCCAAGACCAATACCCGCGAATCTAGGCAAGAAGCTCTCTCTCAGACAACCAATCTC 6540  
 6541 ATGGGCCCCATTCAGGACATCTCTGAGTACATCTTTCATCTCTCTCTCTCTCTCTCTCT 6600  
 6541 ATGGGCCCCATTCAGGACATCTCTGAGTACATCTTTCATCTCTCTCTCTCTCTCTCTCT 6600  
 6601 TGAAGAACCTTACAGTCTCAGGTTCTCGAATCTCTACAGTGCACCTCTGACAGGCT 6660  
 6601 TGAAGAACCTTACAGTCTCAGGTTCTCGAATCTCTACAGTGCACCTCTGACAGGCT 6660  
 6661 CACAGAGTGCACCACTCAACATCATAGTGGAGGCACTGAAAGACACAGCAGAGGCATAA 6720  
 6661 CACAGAGTGCACCACTCAACATCATAGTGGAGGCACTGAAAGACACAGCAGAGGCATAA 6720  
 6721 GGTTCGGGAGAGGTTTACCGTGGGCACTCTGTCAACGAAGGCTTGAACCAACCTTAC 6780  
 6721 GGTTCGGGAGAGGTTTACCGTGGGCACTCTGTCAACGAAGGCTTGAACCAACCTTAC 6780  
 6781 GGAATGACTCGTCTTGAACCCCTACACAGTGTCCCAATATGCGTGGAGATGAGTGGGA 6840  
 6781 GGAATGACTCGTCTTGAACCCCTACACAGTGTCCCAATATGCGTGGAGATGAGTGGGA 6840  
 6841 ACGAATGCTGAATCAGGCTTAACTGTGTGTCAGTCTTACGCTTGGAGTGGTCA 6900  
 6841 ACGAATGCTGAATCAGGCTTAACTGTGTGTCAGTCTTACGCTTGGAGTGGTCA 6900  
 6901 TTTTCAGATGATTCATCTAGATGGTGCATGACATGTTGAACTACAAATGAGGAG 6960  
 6901 TTTTCAGATGATTCATCTAGATGGTGCATGACATGTTGAACTACAAATGAGGAG 6960  
 6961 GAAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGTGCACATGCTTTGGGAACGG 7020  
 6961 GAAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGTGCACATGCTTTGGGAACGG 7020  
 7021 AAAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTACGATGATGGAAGACATA 7080  
 7021 AAAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTACGATGATGGAAGACATA 7080  
 7081 CCACGTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTGCAATGCTT 7140  
 7081 CCACGTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTGCAATGCTT 7140  
 7141 TGGAGCCAGCGGGGTGGGCTGTGACAACTGCGGACACCTGGGGTGAACCCAGTCC 7200  
 7141 TGGAGCCAGCGGGGTGGGCTGTGACAACTGCGGACACCTGGGGTGAACCCAGTCC 7200  
 7201 CGAAGGCACTACTTGGCCAGTCTTACAAACCAAGTATCTCAGAGATACCAATCAGAGAACAA 7260  
 7201 CGAAGGCACTACTTGGCCAGTCTTACAAACCAAGTATCTCAGAGATACCAATCAGAGAACAA 7260  
 7261 CACTAATGTTAATGGCCAAATGAGTGTCTTACGCTTACGCTTACGCTTACGCTTACGCTT 7320  
 7261 CACTAATGTTAATGGCCAAATGAGTGTCTTACGCTTACGCTTACGCTTACGCTTACGCTT 7320  
 7321 AGATTCGGAGAGTAATCATCTTTTCAATCCAGAGGAACCAAGCATGCTCTCTGCGAAG 7380  
 7321 AGATTCGGAGAGTAATCATCTTTTCAATCCAGAGGAACCAAGCATGCTCTCTGCGAAG 7380  
 7381 ATCCATCTAACTGGAGTATGTTAGACACCCAGCTTACGCTTCTCTCTCTCTCTCTCTAA 7440  
 7381 ATCCATCTAACTGGAGTATGTTAGACACCCAGCTTACGCTTCTCTCTCTCTCTCTCTAA 7440  
 7441 GCCCTTGTCTCTGGAGGAGTCTCTCCAGTCTTCACTCACTCAGCTTCTCCAGCATC 7500

Db 7441 GCCCTTGTCTCTGGAGGAGTCTCTCCAGCTTCTCAGCTCAATCAGAGTCTCTCCAGCATC 7500  
 Oy 7501 ACCCTGGAGTCTCTGGAGGTTTCTCATAAATGAGGGCTGCACATTCCTGCTTCTGCT 7560  
 Db 7501 ACCCTGGAGTCTCTGGAGGTTTCTCATAAATGAGGGCTGCACATTCCTGCTTCTGCT 7560  
 Oy 7561 TCGAAGTATTCAATACACGCTCAGTATTAAATGAAGTGAATCTTAAGATTGGTTTGGGA 7620  
 Db 7561 TCGAAGTATTCAATACACGCTCAGTATTAAATGAAGTGAATCTTAAGATTGGTTTGGGA 7620  
 Oy 7621 TCAATAGAAAGCATATGACAGCCAAACCAAGATCAAAATTTTGAATGATATGACCAAA 7680  
 Db 7621 TCAATAGAAAGCATATGACAGCCAAACCAAGATCAAAATTTTGAATGATATGACCAAA 7680  
 Oy 7681 ATTTTAAAGTGAAGTCAACCCCAACACCTTCTGCTTTCACCTTAAGTCTCTGCGCCGCAAT 7740  
 Db 7681 ATTTTAAAGTGAAGTCAACCCCAACACCTTCTGCTTTCACCTTAAGTCTCTGCGCCGCAAT 7740  
 Oy 7741 ACTCTAGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795  
 Db 7741 ACTCTAGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795

RESULT 2  
 HSM806805  
 LOCUS Homo sapiens mRNA; cDNA DKFZp686M2451 (from clone DKFZp686M2451);  
 DEFINITION complete cds.  
 ACCESSION BX640731  
 VERSION BX640731.1 GI:34364819  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 8042)  
 AUTHORS Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,  
 Mewes H.W., Well B., Amid C., Osanger A., Fobo G., Han M. and  
 Wiemann S.  
 The German Human cDNA Consortium  
 Direct Submission  
 Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
 Neuherberg, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by DKFZ (German Cancer Research Center,  
 Heidelberg/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 This clone (DKFZp686M2451) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://mips.gsf.de/proj/cDNA/>.  
 Location/Qualifiers  
 1..8042  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="2q35"  
 /clone="DKFZp686M2451"  
 /tissue\_type="human endometrium carcinoma cell line"  
 /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiA + SfiIB"  
 /dev\_stage="adult"  
 1..8042  
 /gene="DKFZp686M2451"  
 <1..7335  
 /gene="DKFZp686M2451"  
 /note="fibronection precursor, elongation of ORF"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA845847.1"

FEATURES  
 source  
 1..8042  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="2q35"  
 /clone="DKFZp686M2451"  
 /tissue\_type="human endometrium carcinoma cell line"  
 /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiA + SfiIB"  
 /dev\_stage="adult"  
 1..8042  
 /gene="DKFZp686M2451"  
 <1..7335  
 /gene="DKFZp686M2451"  
 /note="fibronection precursor, elongation of ORF"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA845847.1"

gene  
 CDS







3541 CCACCTTAAACACCGAGGTGACTGAGACCAACCAATTGTGATCAATGAGACGCTGTCTCCA 3600 Db  
3602 AGAATTGCTTTTAAGCTCGGTGTACGACCAAGCCAGGAGAGAGGACACACAGAGAAGTG 3661 Qy  
3601 AGAATTGCTTTTAAGCTCGGTGTACGACCAAGCCAGGAGAGAGGACACACAGAGAAGTG 3660 Db  
3662 ACTTCAGACTCAGGAAGCATCGTTGTGTGTCGGCTGTGACTCCAGAGTAGAATAGCTCTAC 3721 Qy  
3661 ACTTCAGACTCAGGAAGCATCGTTGTGTGTCGGCTGTGACTCCAGAGTAGAATAGCTCTAC 3720 Db  
3722 ACCATCCAAAGTCTCAGAGATGGACAGAGAAAGAGATGGCCAAATTGTAAACAAAGTGGTG 3781 Qy  
3721 ACCATCCAAAGTCTCAGAGATGGACAGAGAAAGAGATGGCCAAATTGTAAACAAAGTGGTG 3780 Db  
3782 ACACCATTTGTCTCCACCAACAAACCTTGATCTGAGGCAAAACCTTGACACTGAGAGTCTC 3841 Qy  
3781 ACACCATTTGTCTCCACCAACAAACCTTGATCTGAGGCAAAACCTTGACACTGAGAGTCTC 3840 Db  
3842 ACAGTCTCTCTGGAGAGAGACACCCACAGACATTAAGTGTATAGATTTACCAACAACC 3901 Qy  
3841 ACAGTCTCTCTGGAGAGAGACACCCACAGACATTAAGTGTATAGATTTACCAACAACC 3900 Db  
3902 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAAGTGGTCCATGCTGATFACAGAGTCC 3961 Qy  
3901 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAAGTGGTCCATGCTGATFACAGAGTCC 3960 Db  
3962 TGCATCTTTGATTAACCTGAGTCCCGCTGGAGTACATGTCAGTGTACACTGTCAAG 4021 Qy  
3961 TGCATCTTTGATTAACCTGAGTCCCGCTGGAGTACATGTCAGTGTACACTGTCAAG 4020 Db  
4022 GATGACAAAGGAAGTGTCCCTATCTCTGATACCAATFATCCAGAGTGTTCCTCCCTCACT 4081 Qy  
4021 GATGACAAAGGAAGTGTCCCTATCTCTGATACCAATFATCCAGAGTGTTCCTCCCTCACT 4080 Db  
4082 GACCTGCGATTCACCAACATTTGGTCCAGACACCAAGCTGTCACCTGGCTCCACCCCA 4141 Qy  
4081 GACCTGCGATTCACCAACATTTGGTCCAGACACCAAGCTGTCACCTGGCTCCACCCCA 4140 Db  
4142 TCCATTTGATTTAAACCAACTTCTCGTGTGGTACTCACCTGTGAAATAAGAGAAAGATGT 4201 Qy  
4141 TCCATTTGATTTAAACCAACTTCTCGTGTGGTACTCACCTGTGAAATAAGAGAAAGATGT 4200 Db  
4202 GCAGAGTGTCAATTTCTCTTCAGACAAATGAGTGTCTTAAACAAATCTCTGCTGGT 4261 Qy  
4201 GCAGAGTGTCAATTTCTCTTCAGACAAATGAGTGTCTTAAACAAATCTCTGCTGGT 4260 Db  
4262 ACAGAAATATGATGAGTGTCTCCAGTGTCTACGAAACAAATGAGACACACCTCTTAGA 4321 Qy  
4261 ACAGAAATATGATGAGTGTCTCCAGTGTCTACGAAACAAATGAGACACACCTCTTAGA 4320 Db  
4322 GGAAGACAGAAACAGGTCTTGATTTCCCAACTGGGATGATCTTTCTGATATTAAGTCC 4381 Qy  
4321 GGAAGACAGAAACAGGTCTTGATTTCCCAACTGGGATGATCTTTCTGATATTAAGTCC 4380 Db  
4382 AACTCTTTTACTGTGACATGGATTTGCTCTCAGACCAACCATCACTGGCTACAGGATCCG 4441 Qy  
4381 AACTCTTTTACTGTGACATGGATTTGCTCTCAGACCAACCATCACTGGCTACAGGATCCG 4440 Db  
4442 CATCATCCGAGCACTTCAGTGGAGACTCCGAGAGATTCGGGTCCCACTCTCGGAAT 4501 Qy  
4441 CATCATCCGAGCACTTCAGTGGAGACTCCGAGAGATTCGGGTCCCACTCTCGGAAT 4500 Db  
4502 TCCATCAGCTTCACCAACTCACTCCAGGACAGAGATGTCGTCAGCATCGTCTCTT 4561 Qy  
4501 TCCATCAGCTTCACCAACTCACTCCAGGACAGAGATGTCGTCAGCATCGTCTCTT 4560 Db  
4562 AATGGCAGAGAGAAAGTCCCTTTATTTGATTTGGCCAAACATCAACATTTCTGATTTCCG 4621 Qy  
4561 AATGGCAGAGAGAAAGTCCCTTTATTTGATTTGGCCAAACATCAACATTTCTGATTTCCG 4620 Db  
4622 AGGACCTGGAAGTGTGTCGACACCCCAACAGCTTACTGATCAGTGGGATGCTCT 4681 Qy  
4621 AGGACCTGGAAGTGTGTCGACACCCCAACAGCTTACTGATCAGTGGGATGCTCT 4680 Db

4682 GCTGTCCACAGTGGATATTATACAGGATCACTTACGAGAAACAGGAGGAAATAGCCCTGTC 4741 Qy  
4681 GCTGTCCACAGTGGATATTATACAGGATCACTTACGAGAAACAGGAGGAAATAGCCCTGTC 4740 Db  
4742 CAGGAGTTCACATGCTCGCTGGAGCAAGTCTACAGCTACCATCAGCGGCTTAAACCTGGA 4801 Qy  
4741 CAGGAGTTCACATGCTCGCTGGAGCAAGTCTACAGCTACCATCAGCGGCTTAAACCTGGA 4800 Db  
4802 GTTGATTTATPACCATCACTGTGTATGCTCACTGTCGCTGGAGACAGCCCCCGCAAGCAGC 4861 Qy  
4801 GTTGATTTATPACCATCACTGTGTATGCTCACTGTCGCTGGAGACAGCCCCCGCAAGCAGC 4860 Db  
4862 AAGCCATTTCCATTAATTTACCGAACAGAAATGACAAACCATCCAGATGCAAGTGACC 4921 Qy  
4861 AAGCCATTTCCATTAATTTACCGAACAGAAATGACAAACCATCCAGATGCAAGTGACC 4920 Db  
4922 GATGTTTCAGGACAAACAGCATTAAGTGTCAAGTGGCTTCAAGTTCGCTTCTTACTGCT 4981 Qy  
4921 GATGTTTCAGGACAAACAGCATTAAGTGTCAAGTGGCTTCAAGTTCGCTTCTTACTGCT 4980 Db  
4982 TACAGAGTAAACCAACCTCCCAAAATTTGACAGGACCAACCAAACTTAAACTCAGCT 5041 Qy  
4981 TACAGAGTAAACCAACCTCCCAAAATTTGACAGGACCAACCAAACTTAAACTCAGCT 5040 Db  
5042 CCAGATCAACCAAGAAATGACTATTGAAAGCTTTCAGGCTTCAGCCCAACAGTGGAGTATGCT 5101 Qy  
5041 CCAGATCAACCAAGAAATGACTATTGAAAGCTTTCAGGCTTCAGCCCAACAGTGGAGTATGCT 5100 Db  
5102 GTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAGACTGCAGTACCAAC 5161 Qy  
5101 GTCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTCAGACTGCAGTACCAAC 5160 Db  
5162 ATTGATCGGCTTAAAGGACTGTCATTTCACTGATGTGGATGTCGATTTCCATCAAAATGCT 5221 Qy  
5161 ATTGATCGGCTTAAAGGACTGTCATTTCACTGATGTGGATGTCGATTTCCATCAAAATGCT 5220 Db  
5222 TGGAAGGCGCCACAGGGGCAAGTTTCCAGGTAACAGGTGACCTCTCAGAGCTCAGGAT 5281 Qy  
5221 TGGAAGGCGCCACAGGGGCAAGTTTCCAGGTAACAGGTGACCTCTCAGAGCTCAGGAT 5280 Db  
5282 GGAATCCATGAGTATTTCCCTGCACCTGATGTTGAAGAGACACTGACAGCTGCAAGGC 5341 Qy  
5281 GGAATCCATGAGTATTTCCCTGCACCTGATGTTGAAGAGACACTGACAGCTGCAAGGC 5340 Db  
5342 CTCAGACCGGGTCTCAGTACACAGTCACTGATGTGGTTCCTTGCACGATGATATGAGAGC 5401 Qy  
5341 CTCAGACCGGGTCTCAGTACACAGTCACTGATGTGGTTCCTTGCACGATGATATGAGAGC 5400 Db  
5402 CAGCCCTGATTTGGACCCAGTCCACAGCTATTTCTGACACCAACTGACCTGAAGTTCACT 5461 Qy  
5401 CAGCCCTGATTTGGACCCAGTCCACAGCTATTTCTGACACCAACTGACCTGAAGTTCACT 5460 Db  
5462 CAGGTCAACCCCAACAGCTGAGCGCCAGTGGACACCAACCCCAATTTTCACTGCACTGGA 5521 Qy  
5461 CAGGTCAACCCCAACAGCTGAGCGCCAGTGGACACCAACCCCAATTTTCACTGCACTGGA 5520 Db  
5522 TATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTGTCT 5581 Qy  
5521 TATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTGTCT 5580 Db  
5582 CCTGACAGCTTCATCCCTGTTGTTATCAGGACTTATGTTGGCCCAACCAATATGAGTGA 5641 Qy  
5581 CCTGACAGCTTCATCCCTGTTGTTATCAGGACTTATGTTGGCCCAACCAATATGAGTGA 5640 Db  
5642 GTCTATGCTTTAAGGACACTTTTGAACAGCAGACAGCTCAGGGAGTTGTCACCACTCTG 5701 Qy  
5641 GTCTATGCTTTAAGGACACTTTTGAACAGCAGACAGCTCAGGGAGTTGTCACCACTCTG 5700 Db  
5702 GAGAAATGTCAGGCCCAACCAAGAGGGTCTGTTGACAGATGCTACTGAGACCAACCATCACC 5761 Qy  
5701 GAGAAATGTCAGGCCCAACCAAGAGGGTCTGTTGACAGATGCTACTGAGACCAACCATCACC 5760 Db

5762 ATTAGCTGGAGAACCAAGACTGAGACGATCAGCTGGCTCCAAAGTTGATGCGCTTCAGCC 5821  
5761 ATTAGCTGGAGAACCAAGACTGAGACGATCAGCTGGCTCCAAAGTTGATGCGCTTCAGCC 5820  
5822 AATGGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACACCATCACT 5881  
5821 AATGGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACACCATCACT 5880  
5882 GGGTTACAAACCGGCACTGACTACAGATCTACCTGTACACCTTGAATGACAACTGCTGG 5941  
5881 GGGTTACAAACCGGCACTGACTACAGATCTACCTGTACACCTTGAATGACAACTGCTGG 5940  
5942 AGCTCCCTGTGTATCGACGCTCCATCGCAATGATGCAACCATCCAACTCGCTTC 6001  
5941 AGCTCCCTGTGTATCGACGCTCCATCGCAATGATGCAACCATCCAACTCGCTTC 6000  
6002 CTGGCCACACACCCCAATTCCTGTGCTGTATCATGCGACGCCGCGCTGCGAGGATACC 6061  
6001 CTGGCCACACACCCCAATTCCTGTGCTGTATCATGCGACGCCGCGCTGCGAGGATACC 6060  
6062 GGCTACATCATCAAGTATGAGAGCCCTGGGTCTCTCCCGAGAGAGTGTCTCCCTCGGCC 6121  
6061 GGCTACATCATCAAGTATGAGAGCCCTGGGTCTCTCCCGAGAGAGTGTCTCCCTCGGCC 6120  
6122 CGCCCTGGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAACCGAATATACATTT 6181  
6121 CGCCCTGGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAACCGAATATACATTT 6180  
6182 TATGTCAATTCCTGCTGAAGATATCAGAGAGCGAGCCCTGANTGGAAGGAAAGACA 6241  
6181 TATGTCAATTCCTGCTGAAGATATCAGAGAGCGAGCCCTGANTGGAAGGAAAGACA 6240  
6242 GAGAGCTTCCCAATCTGTGTAACCTTCCACACCCCAATCTTCATGACAGAGATCTTG 6301  
6241 GAGAGCTTCCCAATCTGTGTAACCTTCCACACCCCAATCTTCATGACAGAGATCTTG 6300  
6302 GATGTTCTTCCACAGTCAAAAGAGCCCTTCTGTCACCCACCCCTGGGTATCACACTGA 6361  
6301 GATGTTCTTCCACAGTCAAAAGAGCCCTTCTGTCACCCACCCCTGGGTATCACACTGA 6360  
6362 AATGTTATCAGTTCTCTGGCACTTCTGTGCAAGAACCCAGTGTGGGCAACAATATGTC 6421  
6361 AATGTTATCAGTTCTCTGGCACTTCTGTGCAAGAACCCAGTGTGGGCAACAATATGTC 6420  
6422 TTGAGGACATGTTTATAGGCGGACACACCGCCACCAACCGCCACCCCTAATAGGAT 6481  
6421 TTGAGGACATGTTTATAGGCGGACACACCGCCACCAACCGCCACCCCTAATAGGAT 6480  
6482 AGCCCAAGACCAATACCCCGCAATGTAGGACAGAAAGCTCTCTCAGACAAACCTCTCA 6541  
6481 AGCCCAAGACCAATACCCCGCAATGTAGGACAGAAAGCTCTCTCAGACAAACCTCTCA 6540  
6542 TGGGCCCATTCAGACACTTCTGATACATCATTTCAATGTCATCTGCTGGCACTGAT 6601  
6541 TGGGCCCATTCAGACACTTCTGATACATCATTTCAATGTCATCTGCTGGCACTGAT 6600  
6602 GAAGAACCCTTACAGTTTACGGGTCTCTGGAACCTTCTACAGTGCACCTCTGACGGCTC 6661  
6601 GAAGAACCCTTACAGTTTACGGGTCTCTGGAACCTTCTACAGTGCACCTCTGACGGCTC 6660  
6662 ACCAGAGTGCACCTTACACATCATGTAGGAGGCACTGAAGACCCAGAGGATTAAG 6721  
6661 ACCAGAGTGCACCTTACACATCATGTAGGAGGCACTGAAGACCCAGAGGATTAAG 6720  
6722 GTTCGGAGAGAGTGTGTACCGTGGGCACTCTGTCAACGAAGGCTTGAACCAACCTACG 6781  
6721 GTTCGGAGAGAGTGTGTACCGTGGGCACTCTGTCAACGAAGGCTTGAACCAACCTACG 6780  
6782 GATGACTCGTCTTTGACCCCTTACAGTGTCCATTTATGCGTTGGAGATGAGTGGGAA 6841  
6781 GATGACTCGTCTTTGACCCCTTACAGTGTCCATTTATGCGTTGGAGATGAGTGGGAA 6840  
6842 CGAATGTCATCAAGCTTTAAACTGTTGTGCGCAGTGTAGGCTTTGGAAGTGTCTAT 6901

6841 CGAATGTCATCAAGCTTTAAACTGTTGTGCGCAGTGTAGGCTTTGGAAGTGTCTAT 6900  
6902 TTCAAGTGTGATTCATCTAGATGTCATGACAAATGGTGTGAACTACAAAGATTTGGAGAG 6961  
6901 TTCAAGTGTGATTCATCTAGATGTCATGACAAATGGTGTGAACTACAAAGATTTGGAGAG 6960  
6962 AAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATCTCTTGGGAACCGA 7021  
6961 AAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATCTCTTGGGAACCGA 7020  
7022 AAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTACGATGATGGAGACATAC 7081  
7021 AAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTATGATGATGGAGACATAC 7080  
7082 CACGTAGGAGAACAGTGGCAGAGGAATATCTCGTGCCATTTGCTCCTGCACTGCTTT 7141  
7081 CACGTAGGAGAACAGTGGCAGAGGAATATCTCGTGCCATTTGCTCCTGCACTGCTTT 7140  
7142 GGAGGCCAGCGGCTGGCGCTGTGACAACTGGCGCAGACCTGGGGGTGAACCCAGTCCC 7201  
7141 GGAGGCCAGCGGCTGGCGCTGTGACAACTGGCGCAGACCTGGGGGTGAACCCAGTCCC 7200  
7202 GAAGGCACTACTGGCCAGTCTTACACCACTGATCTCAGAGATACCATCAGAGAACAAAC 7261  
7201 GAAGGCACTACTGGCCAGTCTTACACCACTGATCTCAGAGATACCATCAGAGAACAAAC 7260  
7262 ACTAATGTTAATTTGCCCAATGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAA 7321  
7261 ACTAATGTTAATTTGCCCAATGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAA 7320  
7322 GATTTCCGAGAGTAATCATCTTCCATCCAGAGAACAGCATGTCTCTCTGCCAAGA 7381  
7321 GATTTCCGAGAGTAATCATCTTCCATCCAGAGAACAGCATGTCTCTCTGCCAAGA 7380  
7382 TCCATCTAAACTGGAGTGTAGTGTAGCAGACCCAGCTTTCTCTCTTTCTTTAAG 7441  
7381 TCCATCTAAACTGGAGTGTAGTGTAGCAGACCCAGCTTTCTCTCTTTCTTTAAG 7440  
7442 CCCTTCTCTGGAGGAGTCTTCCAGCTTCACTCACTCAGCTTCTTCCAGGATCA 7501  
7441 CCCTTCTCTGGAGGAGTCTTCCAGCTTCACTCACTCAGCTTCTTCCAGGATCA 7500  
7502 CCCTGGAGTCTTCTGGAGGTTTCTCATAAATGAGGCTGACATTTGCTTCTGCTT 7561  
7501 CCCTGGAGTCTTCTGGAGGTTTCTCATAAATGAGGCTGACATTTGCTTCTGCTT 7560  
7562 CGAAGTATTCATACCGCTCAGTATTTTAAATGAGTGTCTAAAGATTTGGTTGGGAT 7621  
7561 CGAAGTATTCATACCGCTCAGTATTTTAAATGAGTGTCTAAAGATTTGGTTGGGAT 7620  
7622 CAATAGGAAAGCATATCGACCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAA 7681  
7621 CAATAGGAAAGCATATCGACCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAA 7680  
7682 TTTTAAAGTGGAGAAAGTCAACCAACCACTTCTGCTTTCACTTAAAGTGTCTGCGCCGCAATA 7741  
7681 TTTTAAAGTGGAGAAAGTCAACCAACCACTTCTGCTTTCACTTAAAGTGTCTGCGCCGCAATA 7740  
7742 CTGTAGGAAAGCATATCTGTTACTGTATTTTAAATGATATTTTAAATGATATCCAGTACT 7795  
7741 CTGTAGGAAAGCATATCTGTTACTGTATTTTAAATGATATTTTAAATGATATCCAGTACT 7794

## RESULT 3

LOCUS HSM806170 8030 bp mRNA linear PRI 17-JUN-2003  
DEFINITION Homo sapiens mRNA; cDNA DKFp6860016 (from clone DKFp6860016).  
ACCESSION BX538017  
VERSION BX538017.1 GI:31874108  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens















302 CTGGCGTCTGTGCTGGGGAAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGG 361  
301 CTGGCGTCTGTGCTGGGGAAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGG 360  
362 CAGGCTCAGCAAAATGGTTCAGCCCGGAGTCCCGGTGGCTGTGAGTCAAAAGCAAGCCCGGT 421  
361 CAGGCTCAGCAAAATGGTTCAGCCCGGAGTCCCGGTGGCTGTGAGTCAAAAGCAAGCCCGGT 420  
422 TGTATTGACAAATGGGAAAACATATACAGATAAATCAACAGTGGGAGCGGACCTACCTAGGC 481  
421 TGTATTGACAAATGGGAAAACATATACAGATAAATCAACAGTGGGAGCGGACCTACCTAGGC 480  
482 AATGGCTGTGTTGTTACTTGTATGAGGAGGCGGAGGTTTAACTCGGAAAAGTAAACCT 541  
481 AATGGCTGTGTTGTTACTTGTATGAGGAGGCGGAGGTTTAACTCGGAAAAGTAAACCT 540  
542 GAAGCTGAAGAGACTTGTCTTGAAGAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 601  
541 GAAGCTGAAGAGACTTGTCTTGAAGAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 600  
602 TATGAGCGTCTTAAGACTCCATGATCTGGGAGTGTACCTGGGCTGGGCTGGGAGGG 661  
601 TATGAGCGTCTTAAGACTCCATGATCTGGGAGTGTACCTGGGCTGGGCTGGGAGGG 660  
662 AGAATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCACTCTTCAAGATTGGT 721  
661 AGAATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCACTCTTCAAGATTGGT 720  
722 GACACTGGAGGAGACCATGAGACTGGTGGTATCATGTAGAGTGTGTCTTGGT 781  
721 GACACTGGAGGAGACCATGAGACTGGTGGTATCATGTAGAGTGTGTCTTGGT 780  
782 AATGGAAGAGGAGATGGACCTCAAGACCCATAGCTGAGAAGTGTCTTCAATGCTGCT 841  
781 AATGGAAGAGGAGATGGACCTCAAGACCCATAGCTGAGAAGTGTCTTCAATGCTGCT 840  
842 GGGACTCTTATGTGTCGAGGAAAGCTGGGAGAGCCCTACCAAGGCTGGATGTGTA 901  
841 GGGACTCTTATGTGTCGAGGAAAGCTGGGAGAGCCCTACCAAGGCTGGATGTGTA 900  
902 GATTGTACTTTCCTGGGAGAGGACGGGACGGATCACTTGCACTTCTAGAAAATAGATGC 961  
901 GATTGTACTTTCCTGGGAGAGGACGGGACGGATCACTTGCACTTCTAGAAAATAGATGC 960  
962 AACGATCAGAGACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGATAT 1021  
961 AACGATCAGAGACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGATAT 1020  
1022 CGAGGAAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGTGGAAAGTGTGAG 1081  
1021 CGAGGAAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGTGGAAAGTGTGAG 1080  
1082 AGGCAACCTCTCTGTGAGACCAATCGAGGGATCTGGCCCTTCAACCGATGTTGTTGCA 1141  
1081 AGGCAACCTCTCTGTGAGACCAATCGAGGGATCTGGCCCTTCAACCGATGTTGTTGCA 1140  
1142 GCTGTTTACCAACCGGAGCTCACCCCGGCTCTTCCCTTGGCCACTGTCTCAGAC 1201  
1141 GCTGTTTACCAACCGGAGCTCACCCCGGCTCTTCCCTTGGCCACTGTCTCAGAC 1200  
1202 AGTGGTGTGTTACTCTGTGGGATCGAGTGGCTGAAGACCAAGGAAATAGCAATG 1261  
1201 AGTGGTGTGTTACTCTGTGGGATCGAGTGGCTGAAGACCAAGGAAATAGCAATG 1260  
1262 CTTTGGACGTCCTGGGCAACGAGTCACTGGCCAGAGACAGCTGTACCCAGACTTAC 1321  
1261 CTTTGGACGTCCTGGGCAACGAGTCACTGGCCAGAGACAGCTGTACCCAGACTTAC 1320  
1322 GGTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCATTACCTTACAAATGGCAGGCGTTTC 1381  
1321 GGTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCATTACCTTACAAATGGCAGGCGTTTC 1380  
1382 TACTCTGCAACCAAGAGGGGAGAGGAGCGGACATCTTTGGTGCAGCAACACTTCGAAAT 1441

1381 TACTCTGCAACCAAGAGGGGCGACAGGACGGACATCTTTGGTGCAGCAACACTTCGAAT 1440  
1442 TATGAGCAGGACCAAGAAATATCTTTCTGCACAGACCAACTGTGTTGGTTCAGACTCGA 1501  
1441 TATGAGCAGGACCAAGAAATATCTTTCTGCACAGACCAACTGTGTTGGTTCAGACTCGA 1500  
1502 GGAGGAAATTCGAATGCTGTGCTGTGCACTTCCCTTCTATATCAACCAACCAATTAC 1561  
1501 GGAGGAAATTCGAATGCTGTGCTGTGCACTTCCCTTCTATATCAACCAATTAC 1560  
1562 ACTGATTGCACTTCTGAGGGCAGAGAGACAAATGAAGTGGTGTGGGACCAACAGAAC 1621  
1561 ACTGATTGCACTTCTGAGGGCAGAGAGACAAATGAAGTGGTGTGGGACCAACAGAAC 1620  
1622 TATGATCCGACCAAGAAATTTGGTTCCTGCCCCATGCTGCCCGCCAGAGGAAATCTGCACA 1681  
1621 TATGATCCGACCAAGAAATTTGGTTCCTGCCCCATGCTGCCCGCCAGAGGAAATCTGCACA 1680  
1682 ACCAATGAAGGGGTCACTGATCCGCAATTCAGATCAGTGGGATGAAGCAGCATGACATGGGT 1741  
1681 ACCAATGAAGGGGTCACTGATCCGCAATTCAGATCAGTGGGATGAAGCAGCATGACATGGGT 1740  
1742 CACATGATGAGGTGCACTGTGTTGGAAATGGTGTGGGAAATGGACATGCAATGCTCTAC 1801  
1741 CACATGATGAGGTGCACTGTGTTGGAAATGGTGTGGGAAATGGACATGCAATGCTCTAC 1800  
1802 TCGCAGCTTCGAGATCAGTGTGTTGATGACATCACTTACAAATGTAAGACACACATTC 1861  
1801 TCGCAGCTTCGAGATCAGTGTGTTGATGACATCACTTACAAATGTAAGACACACATTC 1860  
1862 CACAAGCGTCAATGAAGAGGGGACATGCTGAACTGTATCACTTCTGCTCAGGGTTCGGGGC 1921  
1861 CACAAGCGTCAATGAAGAGGGGACATGCTGAACTGTATCACTTCTGCTCAGGGTTCGGGGC 1920  
1922 AGGTGGAAGTGTGATCCGCTCGACCAATGCGAGATTCAGAGACTGGGAGCTTTTATCAA 1981  
1921 AGGTGGAAGTGTGATCCGCTCGACCAATGCGAGATTCAGAGACTGGGAGCTTTTATCAA 1980  
1982 ATTGAGATTCATGGGAGAAATGTCATGCTGTGATGATCCAGATCCTACTGCTATGGC 2041  
1981 ATTGAGATTCATGGGAGAAATGTCATGCTGTGATGATCCAGATCCTACTGCTATGGC 2040  
2042 GGTGGCATTTGGGAGTGGCATTTGCCAACCCTTACAGACCTTACCAAGCTCAAGTGGCT 2101  
2041 GGTGGCATTTGGGAGTGGCATTTGCCAACCCTTACAGACCTTACCAAGCTCAAGTGGCT 2100  
2102 GTCCAAAGTATTTATCACTGAGACTCCGAGTCCAGCCCAACTCCCAACCCCTCCAGTGGAAAT 2161  
2101 GTCCAAAGTATTTATCACTGAGACTCCGAGTCCAGCCCAACTCCCAACCCCTCCAGTGGAAAT 2160  
2162 GCACCAAGCCATCTCACTTTCCAAAGTACATTTCCAGGTGGAGACCTTAAATTTCTGTA 2221  
2161 GCACCAAGCCATCTCACTTTCCAAAGTACATTTCCAGGTGGAGACCTTAAATTTCTGTA 2220  
2222 GGCCTTTGGAAGGAGTACCATACAGGCGCACTTAACTCCTACACCACTCAAGGCGCTG 2281  
2221 GGCCTTTGGAAGGAGTACCATACAGGCGCACTTAACTCCTACACCACTCAAGGCGCTG 2280  
2282 AAGCCTGGTGTGATACGAGGGCGAGCTCATGAGCATCCAGCAGTACGGCCACCAAGAA 2341  
2281 AAGCCTGGTGTGATACGAGGGCGAGCTCATGAGCATCCAGCAGTACGGCCACCAAGAA 2340  
2342 GTGACTCGTCTTGAATTCACCAACCAAGCAGCAGCAGCAGTGTGAGCAGCAACACCGTGT 2401  
2341 GTGACTCGTCTTGAATTCACCAACCAAGCAGCAGCAGCAGTGTGAGCAGCAACACCGTGT 2400  
2402 ACAGGAGAGCAGTCTCCCTTTTCTCTTGTGGCCACTTCTGAAATCTGTGACCGAAATC 2461  
2401 ACAGGAGAGCAGTCTCCCTTTTCTCTTGTGGCCACTTCTGAAATCTGTGACCGAAATC 2460  
2462 ACAGGAGAGTCTTGTGCTCTCTCTGCTCTGAGTTCAGCTTCCGACACCGTGTGGGATCCGG 2521

Db 2461 ACAGCAGTAGCTTTGTGGTCTCTGGGTCTCAGCTTCGAGACCCGTGTGGGATTCGG 2520  
Qy 2522 GTGGAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGATCTTCCAAAGCACA 2581  
Db 2521 GTGGAAATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGATCTTCCAGGACA 2580  
Qy 2582 GCCACTCTGTGGAACATCCCTGACCTGCTTCCTGGCCGGAATAATGATGATGCTAT 2641  
Db 2581 GCCACTCTGTGGAACATCCCTGACCTGCTTCCTGGCCGGAATAATGATGATGCTAT 2640  
Qy 2642 CAGATATCTGAGGATGGGAGCAGAGTTGATCTGTCTACTTCCACAAAACAAAGGCGCT 2701  
Db 2641 CAGATATCTGAGGATGGGAGCAGAGTTGATCTGTCTACTTCCACAAAACAAAGGCGCT 2700  
Qy 2702 GATGCCCTCTCTGACCCGACGTGTGGACCAAGTTGATGACACCTCAATGTTGTGCTGG 2761  
Db 2701 GATGCCCTCTCTGACCCGACGTGTGGACCAAGTTGATGACACCTCAATGTTGTGCTGG 2760  
Qy 2762 AGCAGACCCAGGCTCCCATCAGAGGTGACAGATAGTCTATTGCGCCATCAGTAGAAGT 2821  
Db 2761 AGCAGACCCAGGCTCCCATCAGAGGTGACAGATAGTCTATTGCGCCATCAGTAGAAGT 2820  
Qy 2822 AGCAGCAGAACTCAACTCTGAACTGCAAACTCCGTCACCTCAGTGACTTGCAA 2881  
Db 2821 AGCAGCAGAACTCAACTCTGAACTGCAAACTCCGTCACCTCAGTGACTTGCAA 2880  
Qy 2882 CTTGGTGTCTGATATACATCACTATCTATGCTGTGGAAGAAATCAAGAACTACACCT 2941  
Db 2881 CTTGGTGTCTGATATACATCACTATCTATGCTGTGGAAGAAATCAAGAACTACACCT 2940  
Qy 2942 GTTGTCAATCAAGAAACCACTGGCACCCAGCTCAGATACAGTGCCCTCTCCAGG 3001  
Db 2941 GTTGTCAATCAAGAAACCACTGGCACCCAGCTCAGATACAGTGCCCTCTCCAGG 3000  
Qy 3002 GACCTGAGTTTGTGGAAGTACAGAGTGAAGTCAACATCATGTGGACACCGCTGAG 3061  
Db 3001 GACCTGAGTTTGTGGAAGTACAGAGTGAAGTCAACATCATGTGGACACCGCTGAG 3060  
Qy 3062 AGTGCACTGACCGGTACCTGTGGATGTGATCCCGTCAACCTGCTGGGAGCAGCGG 3121  
Db 3061 AGTGCACTGACCGGTACCTGTGGATGTGATCCCGTCAACCTGCTGGGAGCAGCGG 3120  
Qy 3122 CAGAGGCTGCCATCAGCAGAAACCTTTGCAAGTCAACCGGCTGTCCCTGGGCTC 3181  
Db 3121 CAGAGGCTGCCATCAGCAGAAACCTTTGCAAGTCAACCGGCTGTCCCTGGGCTC 3180  
Qy 3182 ACCTATTACTTCAAGTCTTTGCAAGTCAACCGGCTGTCCCTGGGCTC 3241  
Db 3181 ACCTATTACTTCAAGTCTTTGCAAGTCAACCGGCTGTCCCTGGGCTC 3240  
Qy 3242 CAACAGAACCAACTGGATGCTCCCACTAACTTCAAGTCAAGTCAAGTCAAGTCAAGT 3301  
Db 3241 CAACAGAACCAACTGGATGCTCCCACTAACTTCAAGTCAAGTCAAGTCAAGTCAAGT 3300  
Qy 3302 ACTGCTCTGTGAGTCACTCCACCTCGGCGCAGATACAGGATACCGACTGACCGGTG 3361  
Db 3301 ACTGCTCTGTGAGTCACTCCACCTCGGCGCAGATACAGGATACCGACTGACCGGTG 3360  
Qy 3362 GGCCTTACCCGAGGAGCAGCCAGGAGTCAAGTGGTCCCTCTGTCTCAAGTAC 3421  
Db 3361 GGCCTTACCCGAGGAGCAGCCAGGAGTCAAGTGGTCCCTCTGTCTCAAGTAC 3420  
Qy 3422 CCATGAGGATCTCGAGCTGCATCTGAGTACCCGATCCCTGCTGGCCATGAAGGC 3481  
Db 3421 CCATGAGGATCTCGAGCTGCATCTGAGTACCCGATCCCTGCTGGCCATGAAGGC 3480  
Qy 3482 AACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCAGCTGGAGCTCTATT 3541  
Db 3481 AACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCAGCTGGAGCTCTATT 3540  
Qy 3542 CCACCTTACACCGAGGTGACTGAGACCCACTTGTGATCAGATGAGACCGCTGCTCCA 3601  
Db 3541 CCACCTTACACCGAGGTGACTGAGACCCACTTGTGATCAGATGAGACCGCTGCTCCA 3600

Qy 3602 AGAATTGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAAGTG 3661  
Db 3601 AGAATTGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAAGTG 3660  
Qy 3662 ACTTCAGACTCAGGAGCATGTTGTGTCTGGCTGTACTCCAGGATAGATAGTCTTAC 3721  
Db 3661 ACTTCAGACTCAGGAGCATGTTGTGTCTGGCTGTACTCCAGGATAGATAGTCTTAC 3720  
Qy 3722 ACCATCCAAAGTCTGAGAGATGGACAGAAAGAGATGGCCAAATTGTAACAAAGTGCTG 3781  
Db 3721 ACCATCCAAAGTCTGAGAGATGGACAGAAAGAGATGGCCAAATTGTAACAAAGTGCTG 3780  
Qy 3782 ACACCATTTCTCCACCAACAACTTGTGATCTGGAGGCAAAACCTTGACACTGGAGTCTC 3841  
Db 3781 ACACCATTTCTCCACCAACAACTTGTGATCTGGAGGCAAAACCTTGACACTGGAGTCTC 3840  
Qy 3842 ACAGTCTCTCTGGAGAGGACCAACCCAGACATTTACTGCTTATAGAAATTTACCACAAAC 3901  
Db 3841 ACAGTCTCTCTGGAGAGGACCAACCCAGACATTTACTGGTTATAGAAATTTACCACAAAC 3900  
Qy 3902 CCTACAAAGCCGACAGGAGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGCTCC 3961  
Db 3901 CCTACAAAGCCGACAGGAGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGCTCC 3960  
Qy 3962 TGCACTTTTGAATAACCTGAGTCCCGCTGGAGTACAAATGTCACTGTTTACACTGTCAAG 4021  
Db 3961 TGCACTTTTGAATAACCTGAGTCCCGCTGGAGTACAAATGTCACTGTTTACACTGTCAAG 4020  
Qy 4022 GATGACAAAGGAGTCCCTATCTCTGATACCATCATCCAGCTGTTCTCTCCCACT 4081  
Db 4021 GATGACAAAGGAGTCCCTATCTCTGATACCATCATCCAGCTGTTCTCTCCCACT 4080  
Qy 4082 GACCTGCGATTCACCAACCAATTTGTCACACACCATGCTGTCACCTGGCTCCACCCCA 4141  
Db 4081 GACCTGCGATTCACCAACCAATTTGTCACACACCATGCTGTCACCTGGCTCCACCCCA 4140  
Qy 4142 TCCATTGATTTAAACCACTTCTGCTGCTTACTCACTGTGAAATAATGAGAGATGTT 4201  
Db 4141 TCCATTGATTTAAACCACTTCTGCTGCTTACTCACTGTGAAATAATGAGAGATGTT 4200  
Qy 4202 CGAGAGTTGTCAAATTTCTTCCATCAGCAATGAGTGTCTTAAACAAATCTCTGCTGCT 4261  
Db 4201 CGAGAGTTGTCAAATTTCTTCCATCAGCAATGAGTGTCTTAAACAAATCTCTGCTGCT 4260  
Qy 4262 ACAGAAATGTAGTGTCTTCCAGTGTCTACGAAACCAATGAGACACCTCTTAGA 4321  
Db 4261 ACAGAAATGTAGTGTCTTCCAGTGTCTACGAAACCAATGAGACACCTCTTAGA 4320  
Qy 4322 GGAAGACAGAAACAGGCTTGGATTTCCCACTGGCATTTCTTCTGATATTAAGTCC 4381  
Db 4321 GGAAGACAGAAACAGGCTTGGATTTCCCACTGGCATTTCTTCTGATATTAAGTCC 4380  
Qy 4382 AACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCACCATCACTGGCTACAGGATCCGC 4441  
Db 4381 AACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCACCATCACTGGCTACAGGATCCGC 4440  
Qy 4442 CATCATCCGAGACCTTCACTGGAGAGCTTCCAGAGATCGGTCGCCACTCTCGAAT 4501  
Db 4441 CATCATCCGAGACCTTCACTGGAGAGCTTCCAGAGATCGGTCGCCACTCTCGAAT 4500  
Qy 4502 TCCATCACTCCACCAACCTCACTCCAGGACAGAGTATGCTGTCAGCATCTGCTCTT 4561  
Db 4501 TCCATCACTCCACCAACCTCACTCCAGGACAGAGTATGCTGTCAGCATCTGCTCTT 4560  
Qy 4562 AATGGCAGAGAGAAAGTCCCTTATTTGATGGCCAAACATCAAGTCTTCTGATTCGG 4621  
Db 4561 AATGGCAGAGAGAAAGTCCCTTATTTGATGGCCAAACATCAAGTCTTCTGATTCGG 4620  
Qy 4622 AGGACCTCGAAGTGTGCTGCGACCCCACTACTGATCAGCTGGAGTCTCT 4681  
Db 4621 AGGACCTCGAAGTGTGCTGCGACCCCACTACTGATCAGCTGGAGTCTCT 4680







2y	122	CCCTTCCCACCCCTCTGGCCCCCACCTTCTTGAGGCGCAACAACCCCGGAGGCATTAGA	181
2b	123	CCCTTCCCACCCCTCTGGCCCCCACCTTCTTGAGGCGCAACAACCCCGGAGGCATTAGA	182
2y	182	AGGGATTTTCCCGCAGGTGCGAAGGGAAGCAAACTTGGTGGCACTTTGCCCTCCGGTG	241
2b	183	AGGGAATTTTCCCGCAGGTGCGAAGGGAAGCAAACTTGGTGGCACTTTGCCCTCCGGTG	242
2y	242	CGGCGCTCTCTCCCCCACCCGCTCTCAACATGCTTAGGGGTCCGGGGCCGGGGTGTGTGTG	301
2b	243	CGGCGCTCTCTCCCCCACCCGCTCTCAACATGCTTAGGGGTCCGGGGCCGGGGTGTGTGTG	302
2y	302	CTGGCCGTCTGTGCTCTGGGGAAGCGGTGCCCCCTCCACGGGAGCTTCGAAGAGCAAGAGG	361
2b	303	CTGGCCGTCTGTGCTCTGGGGAAGCGGTGCCCCCTCCACGGGAGCTTCGAAGAGCAAGAGG	362
2y	362	CAGGCTCAGCAAAATGGTTTCAAGCCCCAGTCCCGGTGGCTGTCAGTCAAAGCAAGCCCGGT	421
2b	363	CAGGCTCAGCAAAATGGTTTCAAGCCCCAGTCCCGGTGGCTGTCAGTCAAAGCAAGCCCGGT	422
2y	422	TGTTATGACAATGGAAAAACAATATCAGATAAATCAACAGTGGAGCGGACCTACTTAGGC	481
2b	423	TGTTATGACAATGGAAAAACAATATCAGATAAATCAACAGTGGAGCGGACCTACTTAGGC	482
2y	482	AATGCGGTGGTTTGTACTTGTATGAGGGAAGCCGAGGTTTAACTCGGAATCAAACT	541
2b	483	AATGCGGTGGTTTGTACTTGTATGAGGGAAGCCGAGGTTTAACTCGGAATCAAACT	542
2y	542	GAAGCTGAAGAGACTTGTCTTTGCAAGTACACTGGGAACACTTACCAGTGGGTGACCT	601
2b	543	GAAGCTGAAGAGACTTGTCTTTGCAAGTACACTGGGAACACTTACCAGTGGGTGACCT	602
2y	602	TATGAGCGTCTTAAAGACTCCATGATCTGGACTGTGTACTGTGCATCGGGCTCGGCGAGG	661
2b	603	TATGAGCGTCTTAAAGACTCCATGATCTGGACTGTGTACTGTGCATCGGGCTCGGCGAGG	662
2y	662	AGAATAAGCTGTACCATCGCAAAACCGCTGCATGAAGGGGTCACTCAAGAATGGT	721
2b	663	AGAATAAGCTGTACCATCGCAAAACCGCTGCATGAAGGGGTCACTCAAGAATGGT	722
2y	722	GACACTGGAGGAGACACATGAGACTGTGTGGTTACATGTTAGAGTGTGTGTCTTGGT	781
2b	723	GACACTGGAGGAGACACATGAGACTGTGTGGTTACATGTTAGAGTGTGTGTCTTGGT	782
2y	782	AATGGAAAAAGAGATGGACTCTGCAAGCCCATAGCTGAGAAGTGTTTGATCATGTGCT	841
2b	783	AATGGAAAAAGAGATGGACTCTGCAAGCCCATAGCTGAGAAGTGTTTGATCATGTGCT	842
2y	842	GGGACTTCTTATGTGTGCGAGAAACGTGGGAGAGCCCTACCAAGCTCGATGATGTA	901
2b	843	GGGACTTCTTATGTGTGCGAGAAACGTGGGAGAGCCCTACCAAGCTCGATGATGTA	902
2y	902	GATTTACTTTCCTGGGAGAGGAGCGGAGCGCATCACTTGCATCTCAAAATAGATGC	961
2b	903	GATTTACTTTCCTGGGAGAGGAGCGGAGCGCATCACTTGCATCTCAAAATAGATGC	962
2y	962	AACGATCAGGACACAAGGACATCTATAGAAATGGAGACACCTGGAGCAAGAAGGATAAT	1021
2b	963	AACGATCAGGACACAAGGACATCTATAGAAATGGAGACACCTGGAGCAAGAAGGATAAT	1022
2y	1022	CGAGGAAACCTGCTCAGTGATCTGCACAGGCAAGCCCGAGGAGTGGAGTGTGAG	1081
2b	1023	CGAGGAAACCTGCTCCTCAGTGATCTGCACAGGCAAGCCCGAGGAGTGGAGTGTGAG	1082
2y	1082	AGGCAACCTCTGTGCAGACACACATCGAGGGGATCTGGCCCTTCAACCGATGTTCTGCA	1141
2b	1083	AGGCAACCTCTGTGCAGACACACATCGAGGGGATCTGGCCCTTCAACCGATGTTCTGCA	1142
2y	1142	GCTGTTTACCAACCGAGCCTCAACCCAGCCTCCTCCTATGCGCACTGTGTCAAGAC	1201
2b	1143	GCTGTTTACCAACCGAGCCTCAACCCAGCCTCCTCCTATGCGCACTGTGTCAAGAC	1202

Qy	1202	AGTGGTGTGGTCTTACTCTCTGTGGGATGCAAGTGGCTGAAGACACACAGGAAATTAAGCAAAATG	1261
Db	1203	AGTGGTGTGGTCTACTCTCTGTGGGATGCAAGTGGCTGAAGACACAGGAAATTAAGCAAAATG	1262
Qy	1262	CTTTGCAAGTGCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAACCCAGACTTAC	1321
Db	1263	CTTTGCAAGTGCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAACCCAGACTTAC	1322
Qy	1322	GGTGGCAACTCAAAATGGAGAGCGCATGTGTCTTAACAATTCACTTAAGTGGCAGGACGTTTC	1381
Db	1323	GGTGGCAACTCAAAATGGAGAGCGCATGTGTCTTAACAATTCACTTAAGTGGCAGGACGTTTC	1382
Qy	1382	TACTCTCTGCACACACAGAGGGCGACAGGACGGAGACATCTTTGGTGCAGACCAACTTCGAAT	1441
Db	1383	TACTCTCTGCACACAGAGGGCGACAGGACGGAGACATCTTTGGTGCAGACCAACTTCGAAT	1442
Qy	1442	TATGAGCAGSACCAGAAAATACTCTTTCTGCACAGACCAACATGTTTGGTTCAGACTCGA	1501
Db	1443	TATGAGCAGSACCAGAAAATACTCTTTCTGCACAGACCAACATGTTTGGTTCAGACTCGA	1502
Qy	1502	GGAGGAAATCCAAATGGTGGCTTGTGGCACTTCCCCTTCCCTATACAAACCAACCAATTTAC	1561
Db	1503	GGAGGAAATCCAAATGGTGGCTTGTGGCACTTCCCCTTCCCTATACAAACCAACCAATTTAC	1562
Qy	1562	ACTGATTGCACCTTCTCAGGGCCAGAAAGAGACAAACATGAAGTGGTGTGGGACCAACAGAGAAC	1621
Db	1563	ACTGATTGCACCTTCTCAGGGCCAGAAAGAGACAAACATGAAGTGGTGTGGGACCAACAGAGAAC	1622
Qy	1622	TATGATGCCGACACAGAAAGTTGGGTTCTGCCCATAGCTGCCACACAGGNAATCTGCACA	1681
Db	1623	TATGATGCCGACACAGAAAGTTGGGTTCTGCCCATAGCTGCCACACAGGNAATCTGCACA	1682
Qy	1682	ACCAATGAAGGGGTCACTGATACCCCATTTGAGATGACAGTGGGATGAAGCAGATGACATGGGT	1741
Db	1683	ACCAATGAAGGGGTCACTGATACCCCATTTGAGATGACAGTGGGATGAAGCAGATGACATGGGT	1742
Qy	1742	CACATGATAGGTGCACGTGTGTTGGGAATGTCGTGGGGAATGSCATGCAATTCCTAC	1801
Db	1743	CACATGATAGGTGCACGTGTGTTGGGAATGTCGTGGGGAATGSCATGCAATTCCTAC	1802
Qy	1802	TCCAGCTTCGAGATCAGATGTCATTTGTTGATGACATCACTTTACAATGTGAACGACACATTC	1861
Db	1803	TCCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTTACAATGTGAACGACACATTC	1862
Qy	1862	CACAAGCGTCATGAAGAGGGGCACATCTGTAACCTGTACATGCTTCGGTCAAGGTGCGGGC	1921
Db	1863	CACAAGCGTCATGAAGAGGGGCACATCTGTAACCTGTACATGCTTCGGTCAAGGTGCGGGC	1922
Qy	1922	AGTGGAAAGTGTGATCCCGTCGACCAATGCGCAGGATTCAGAGACTGGACGTTTTATCAA	1981
Db	1923	AGTGGAAAGTGTGATCCCGTCGACCAATGCGCAGGATTCAGAGACTGGAGCGTTTTATCAA	1982
Qy	1982	ATTGGAGATTCATGGGAGAAAGTATGTGCATGTGTGATGATACCAAGTGCCTACTCTATGGC	2041
Db	1983	ATTGGAGATTCATGGGAGAAAGTATGTGCATGTGTGATGATACCAAGTGCCTACTCTATGGC	2042
Qy	2042	CGTGGCATTTGGGAGTGGCATTCGCCAACCTTTACAGACCTATCCAGCTCAAGTGTCTCT	2101
Db	2043	CGTGGCATTTGGGAGTGGCATTCGCCAACCTTTACAGACCTATCCAGCTCAAGTGTCTCT	2102
Qy	2102	GTCCAAAGTATTTTATCACTGAGACTCCGAGTCAGGCCAACTCCCAACCCCATCCAGTGGAAAT	2161
Db	2103	GTCCAAAGTATTTTATCACTGAGACTCCGAGTCAGGCCAACTCCCAACCCCATCCAGTGGAAAT	2162
Qy	2162	GCACCAAGCGCATCTCACTTTCCAGTACATTTCTAGTGGGAGACCTTAATAATTTCTGTA	2221
Db	2163	GCACCAAGCGCATCTCACTTTCCAGTACATTTCTAGTGGGAGACCTTAATAATTTCTGTA	2222
Qy	2222	GGCGGTTGGAGGAAGCTACCATACAGGCCACTTTAACTCTCTACCACTCAAGAGGCGCTG	2281
Db	2223	GGCGGTTGGAGGAAGCTACCATACAGGCCACTTTAACTCTCTACCACTCAAGAGGCGCTG	2282
Qy	2282	AAGCCTGGTGTGTTATACAGGGGCCAGCTCATCAGCATCCAGCAGTACGGCCACCAAGAA	2341



4502 TCATCACCCTCACCACCTCACTCAGGACAGAGTATGTGTGTCAGATCGTTGCTCTT 4561  
4499 TCATCACCCTCACCACCTCACTCAGGACAGAGTATGTGTGTCAGATCGTTGCTCTT 4558  
4562 AATGGCAGAGAGAAAGTCCCTTATTGATGGCCAAACAATCAACAGTTTCTGATGTCG 4621  
4559 AATGGCAGAGAGAAAGTCCCTTATTGATGGCCAAACAATCAACAGTTTCTGATGTCG 4618  
4622 AGGACCTGGAAGTGTGTGTCGAGACCCACAGCCTACTGATCAGCTGGGATGCTCT 4681  
4619 AGGACCTGGAAGTGTGTGTCGAGACCCACAGCCTACTGATCAGCTGGGATGCTCT 4678  
4682 GCTGTCACTGAGATATTACAGATCACTTACGGAGAAACAGAGGAAATAGCCCTGTC 4741  
4679 GCTGTCACTGAGATATTACAGATCACTTACGGAGAAACAGAGGAAATAGCCCTGTC 4738  
4742 CAGAGTTCACTGCTGGGACAGTCTACAGCTTACCATCAGCGGCTTAAACCTGGA 4801  
4739 CAGAGTTCACTGCTGGGACAGTCTACAGCTTACCATCAGCGGCTTAAACCTGGA 4798  
4802 GTTGATTATACCATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4861  
4799 GTTGATTATACCATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4858  
4862 AAGCCAAATTCATTAATACCGAAACAGAAATTTGACAAACCATCCAGATGCAAGTACC 4921  
4859 AAGCCAAATTCATTAATACCGAAACAGAAATTTGACAAACCATCCAGATGCAAGTACC 4918  
4922 GATGTTTCAGACCAACAGCATTTAGTGTCAAGTGTGCTGCTTCAAGTTCCCTGTTACTG 4981  
4919 GATGTTTCAGACCAACAGCATTTAGTGTCAAGTGTGCTGCTTCAAGTTCCCTGTTACTG 4978  
4982 TAGAGATTAACCACTCCCAAAATGACACAGACCAACCAAAACCTAAATCTGAGGT 5041  
4979 TAGAGATTAACCACTCCCAAAATGACACAGACCAACCAAAACCTAAATCTGAGGT 5038  
5042 CCAGATCAACAGAAATGACTATTGAAGGCTTCAGCCACAGTGGAGTATGTTGTTAGT 5101  
5039 CCAGATCAACAGAAATGACTATTGAAGGCTTCAGCCACAGTGGAGTATGTTGTTAGT 5098  
5102 GTCTATGCTCAGAAATCCAGCGAGAGATCAAGCTCTGTTTCAGACTCAGTAACCAAC 5161  
5099 GTCTATGCTCAGAAATCCAGCGAGAGATCAAGCTCTGTTTCAGACTCAGTAACCAAC 5158  
5162 ATTGATCGCCTAAAGACTGGAATTCATGATGTGGATGTGATTCATCAAAATGCT 5221  
5159 ATTGATCGCCTAAAGACTGGAATTCATGATGTGGATGTGATTCATCAAAATGCT 5218  
5222 TGGGAAAGCCACAGGGGCAAGTTTCCAGGTAAGGGTGACCTACTCGAGCCCTGAGGAT 5281  
5219 TGGGAAAGCCACAGGGGCAAGTTTCCAGGTAAGGGTGACCTACTCGAGCCCTGAGGAT 5278  
5282 GGAATCCATGAGCTATTCCTGCACTGATGTTGAAGAGACACTCGAGAGTGCAGGC 5341  
5279 GGAATCCATGAGCTATTCCTGCACTGATGTTGAAGAGACACTCGAGAGTGCAGGC 5338  
5342 CTCAGACCGGGTCTGAGTACACAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 5401  
5339 CTCAGACCGGGTCTGAGTACACAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 5398  
5402 CAGCCCTGATTTGGAACCCAGTCCAGCTATTTCTGTCACCACTGACCTGAAATTCAT 5461  
5399 CAGCCCTGATTTGGAACCCAGTCCAGCTATTTCTGTCACCACTGACCTGAAATTCAT 5458  
5462 CAGTCAACCCCAACAGCTGAGCCGAGTGCACACCACTGAGCTCACTGGA 5521  
5459 CAGTCAACCCCAACAGCTGAGCCGAGTGCACACCACTGAGCTCACTGGA 5518  
5522 TATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGCT 5581  
5519 TATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGCT 5578

5582 CCTGACAGCTCATCCGTGGTGTATCAGGACTTATGTTGGCCACCAATATGAAGTGT 5641  
5579 CCTGACAGCTCATCCGTGGTGTATCAGGACTTATGTTGGCCACCAATATGAAGTGT 5638  
5642 GTCTATGCTCTTAAGGACACTTTGACAGAGAGAGAGTGTGAGAGTGTGACCACTCTG 5701  
5639 GTCTATGCTCTTAAGGACACTTTGACAGAGAGAGAGTGTGAGAGTGTGACCACTCTG 5698  
5702 GAGAAATGTGACCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACATCAAC 5761  
5699 GAGAAATGTGACCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACATCAAC 5758  
5762 ATTAGCTGGAGAACCAAGACTGAGAGAGATCACTGGCCCTCMAAGTTGATGCGTTCCAGCC 5821  
5759 ATTAGCTGGAGAACCAAGACTGAGAGAGATCACTGGCTTCCAAGTTGATGCGTTCCAGCC 5818  
5822 AATGGCAGAGCTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAGCTACACCATCACT 5881  
5819 AATGGCAGAGCTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAGCTACACCATCACT 5878  
5882 GGCTTAAACCAAGAGACTGAGTCAAGATCTACTGTGACACTTTGAATGACAAATGCTCGG 5941  
5879 GGCTTAAACCAAGAGACTGAGTCAAGATCTACTGTGACACTTTGAATGACAAATGCTCGG 5938  
5942 AGCTCCCTGCTGTCATCGAGCGCTCCACTGCTGCTGATGACACCATCAACCTCGCTTC 6001  
5939 AGCTCCCTGCTGTCATCGAGCGCTCCACTGCTGCTGATGACACCATCAACCTCGCTTC 5998  
6002 CTGGCCACCAACCAATTCCTGCTGCTGATCATGCGAGCGCCACGTCGAGGATTAAC 6061  
5999 CTGGCCACCAACCAATTCCTGCTGCTGATCATGCGAGCGCCACGTCGAGGATTAAC 6058  
6062 GGCTTAACTCATGATGAGAGCGCTGGCTGCTCTCCAGAGAGTGTGCTCGCCG 6121  
6059 GGCTTAACTCATGATGAGAGCGCTGGCTGCTCTCCAGAGAGTGTGCTCGCCG 6118  
6122 CGCCCTGCTGTCACAGAGGCTACTATTACTGCTGCTGAAACCGGAAACCGAATATCAATT 6181  
6119 CGCCCTGCTGTCACAGAGGCTACTATTACTGCTGCTGAAACCGGAAACCGAATATCAATT 6178  
6182 TATGTCATTTGCTGTAAGATTAATCAGAGAGCGGCTGATTTGAGAGAGAGAGAGAGAG 6241  
6179 TATGTCATTTGCTGTAAGATTAATCAGAGAGCGGCTGATTTGAGAGAGAGAGAGAGAG --- 6235  
6242 GACGAGCTTCCCACTGTTAAACCTTCCACACCCCAATCTTATGAGACCAAGATCTTG 6301  
6236 ----- 6235  
6302 GATGTTCTTCCACAGTTCAAAAGACCCCTTTGTCACCAACCCCTGGGTATGACACTGGA 6361  
6236 -----ACAGTTCAAAAGACCCCTTTGTCACCAACCCCTGGGTATGACACTGGA 6283  
6362 AATGTTATTCAGCTTCTGTCAGCTTCTGTCAGCAACCCAGTGTGTCGCAACCAATGATC 6421  
6284 AATGTTATTCAGCTTCTGTCAGCTTCTGTCAGCAACCCAGTGTGTCGCAACCAATGATC 6343  
6422 TTTGAGGAAACATGTTTATGAGCGGACCAACCGCCCAACCGGCAACCCCATTAAGGAT 6481  
6344 TTTGAGGAAACATGTTTATGAGCGGACCAACCGCCCAACCGGCAACCCCATTAAGGAT 6403  
6482 AGGCCAAGACCATACCGCGAATGTAGGACAGAGCTCTCTCTCAGACCAACCATCTCA 6541  
6404 AGGCCAAGACCATACCGCGAATGTAGGACAGAGCTCTCTCTCAGACCAACCATCTCA 6463  
6542 TGGGCCCCCATTTCCAGGACACTTCTGAGTACATCAATTTTCTGTCATCTGTTGGCACTGAT 6601  
6464 TGGGCCCCCATTTCCAGGACACTTCTGAGTACATCAATTTTCTGTCATCTGTTGGCACTGAT 6523  
6602 GAAGAACCTTTACAGTTTCCAGGTTCTGTCAGTCTTCTACAGTGTGCTCTGACAGGCTC 6661  
6524 GAAGAACCTTTACAGTTTCCAGGTTCTGTCAGTCTTCTACAGTGTGCTCTGACAGGCTC 6583  
6662 ACCAGAGTGGCCCTCAACATCATAGTGGAGCCTGAAAGACCAAGAGAGGATTAAG 6721



Db	6584	ACCAGAGTGCACCTCAACAATCATATAGTGGAGCACTGAAGACCAGCAGAGGCATAAG	6643.
Qy	6722	GTTCGGGAAGAGGTTGTTACCGTGGGCAACTCTGCTCAACGAAGGCTTGAAACCAACTACG	6781
Db	6644	GTTCGGGAAGAGGTTGTTACCGTGGGCAACTCTGCTCAACGAAGGCTTGAAACCAACTACG	6703
Qy	6782	GATGACTGCTGTTTGACCCCTACACAGTGTCCCATATATCCGTTGGAGATAGTGGGAA	6841
Db	6704	CATGACTCGTGTGTGACCCCTACACAGTTCCTCAATATGTCGTTGGAGATAGTGGGAA	6763
Qy	6842	CGAATGCTGAATCAGAGCTTTAAACTGTTGTCGAGTGCTTAGGCTTTGGAGTGTGCAT	6901
Db	6764	CGAATGCTGAATCAGAGCTTTAAACTGTTGTCGAGTGCTTAGGCTTTGGAGTGTGCAT	6823
Qy	6902	TTCAGATGTGATTCATCTAGATGGTCCATCATCACAATGGTGTGAACCTACAAGATTGAGAG	6961
Db	6824	TTCAGATGTGATTCATCTAGATGGTCCATCATCACAATGGTGTGAACCTACAAGATTGAGAG	6883
Qy	6962	AAGTGGGACCGTCAGGAGAAATGCGCCAGATGATGAGCTGCACATGCTCTTGGGAACGGA	7021
Db	6884	AAGTGGGACCGTCAGGAGAAATGCGCCAGATGATGAGCTGCACATGCTCTTGGGAACGGA	6943
Qy	7022	AAAGGAGAATTCAGTGTGACCCCTCATGAGGCAACGTGTTACGATGATGGAAGACATAC	7081
Db	6944	AAAGGAGATTCAGTGTGACCCCTCATGAGGCAACGTGTTATGATGATGGAAGACATAC	7003
Qy	7082	CACGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTCTGCACATGCTTT	7141
Db	7004	CACGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTCTGCACATGCTTT	7063
Qy	7142	GGAGCCAGCGGGCTGGCGCTGTGACAACTGCGCGCAGAGCTGGGGGTGAACCCAGTCCC	7201
Db	7064	GGAGCCAGCGGGCTGGCGCTGTGACAACTGCGCGCAGAGCTGGGGGTGAACCCAGTCCC	7123
Qy	7202	GAAGGCATACTGGCCAGTCCCTACAAACAGTATTTCTCAGAGATACCATCAGAGAAACAAAC	7261
Db	7124	GAAGGCATACTGGCCAGTCCCTACAAACAGTATTTCTCAGAGATACCATCAGAGAAACAAAC	7183
Qy	7262	ACTAATGTTAATGCCAATTTAGTGTCTTCATGCTTTTAGATGTACAGGCTGCACAGAA	7321
Db	7184	ACTAATGTTAATGCCAATTTAGTGTCTTCATGCTTTTAGATGTACAGGCTGCACAGAA	7243
Qy	7322	GATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGTCTCTGCGCAAGA	7381
Db	7244	GATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGTCTCTGCGCAAGA	7303
Qy	7382	TCCATCTAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTCTTCTTTCTTTCTTAAG	7441
Db	7304	TCCATCTAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTCTTCTTTCTTTCTTAAG	7363
Qy	7442	CCCTTTGTCTCTGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAAGTCTTCCCAAGCATCA	7501
Db	7364	CCCTTTGTCTCTGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAAGTCTTCCCAAGCATCA	7423
Qy	7502	CCGTGGAGATTTCTGAGGGTTTCTCATAAATGAGGCTGCACATTTGCCGTGTCGCTT	7561
Db	7424	CCGTGGAGATTTCTGAGGGTTTCTCATAAATGAGGCTGCACATTTGCCGTGTCGCTT	7483
Qy	7562	CGAAGTATTCAAATCCGCTCAGTATTTTAAATGAAGTGATTTCTAAGATTTGGTTGGGAT	7621
Db	7484	CGAAGTATTCAAATCCGCTCAGTATTTTAAATGAAGTGATTTCTAAGATTTGGTTGGGAT	7543
Qy	7622	CAATAGGAAGCATATGAGGCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAA	7681
Db	7544	CAATAGGAAGCATATGAGGCCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAA	7603
Qy	7682	TTTTAAAGTAGGAAGTCAACCAACACTTCTGCTTTCACTTAAAGTGTCTGCCCGCAATA	7741
Db	7604	TTTTAAAGTAGGAAGTCAACCAACACTTCTGCTTTCACTTAAAGTGTCTGCCCGCAATA	7663
Qy	7742	CTGTAGGAACCAAGCATGATCTTGTTACTGTGATATTTTAAATATCCCAAGTACT	7795

7664 CWTGAGAACAGCAGATGCTCTGTACTGTATTTAAATATCCACAGTACT 7717

Db

RESULT 6  
 HSM806902  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

7591 bp mRNA linear PRI 30-AUG-2003  
 Homo sapiens mRNA; cDNA DKFP686K139 (from clone DKFP686K139);  
 complete cds.  
 BX649182  
 BX649182.1 GI:34366424

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 7591)  
 Ansong, W., Krieger, S., Regiert, T., Rittmüller, C., Schwager, B.,  
 Mewes, H.W., Weil, B., Amid, C., Oeinger, A., Fobo, G., Han, M. and  
 Wiemann, S.

The German Human cDNA Consortium  
 Direct Submission  
 Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
 Neuberg, GERMANY  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@kfz-heidelberg.de;  
 sequenced by EMBL (European Molecular Biology Laboratories,  
 Heidelberg/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 This clone (DKFP686K139) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://mips.gsf.de/proj/cDNA/>.  
 Location/Qualifiers

1..7951  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="2q35"  
 /clone="DKFP686K139"  
 /tissue\_type="human cervix"  
 /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiIA + SfiIB"  
 /dev\_stage="adult"  
 1..7951  
 /gene="DKFP686K139"  
 <1..3314  
 /gene="DKFP686K139"  
 /note="fibronectin precursor"  
 /codon\_start=3  
 /product="hypothetical protein"  
 /protein\_id="CA546200.1"  
 /db\_xref="GI:34366425"  
 /translation="GPRRLCTGGGEGTGPAGSKGRPAATTSVLVLCIPSPVPPVPPPT  
 LWPWRPQPGIRGDRSRRLREANLVATCLPVRASLPRLNMLRGPGGLLILA  
 VLCLGTAVPTGASKRKQAOQVQSPVAVSQSPGCVNDGKHQVINOQWERYLVG  
 NALVCTCYGSGRGNCFSPSEAEETCFDKYTGNTYVGDYERPKSMINDCTCIGAG  
 RGRISTIANRHEGQGSYKIGDWRRPHTGTGGMRLCEVCLNGKGEWTKCPAECF  
 DHAAGTSYVGETYKPYQGMVMDCTCLGEGSRITCTSRNKGNDTCTSTRIGDT  
 WSKDNRLMLQCITCTNGRGWKCRHSVTQSTSSGSEFTDVAAYVQSPHPQPV  
 PGYCVTDGSAVYSGVMQLKTCNQKMLCTCLNGVSCQETATVOTYQGNSENGCV  
 LPYTGRTFYSTCTSEGRDHLWCSTTSYEQDKYFCTDETLVQTRGNSNGAL  
 CHEPFLNNHYTDCITSEGRDNKWCCTTQNTDAQKFGFCPMAAHEETCTNEGVM  
 YRLGDWDQKDHGMKRCVNGRGWETCIAYSOLRDOCI VDDITVYNDVTFHKRG  
 EBGHMLNCTFGGGRGWKCDPVFVQCQDSEGTGFIQIGDSWEKVGVRVQCVCYGG  
 IEGHMLNCTFGQTPSSSGPVEVFTDTPGNSHPINQNAQPSHISKILLRWPKNVS  
 GRWKEATIPGLHNSYTIKGLKPGVYEGQLISIQQHGQVTRFDTTSTSTPVS  
 TVTGETPSPVLATSEVTEITAFSVFVSWASDTSVGRVSEYELSEEGDPPVLD  
 LPATSVNIPDLPLPKRKIVNYQISEDEGQSLISTQTAPDAPDPTVQDDT  
 SVIVRSVPOAPTGRIVYSPVSEGSSTELNPETANSVTLSDQPGVQVNIITVAV  
 ENQEGTPVYIQETGTGTGTRSDVPSRDLQFVEVTDKVTIMKTPESAVTGVYDVA  
 IPVNLPGHGRQLPISGRNFAEVTLGSPGVYTFKFPVAVSHGRESKPLTAQQTLLDA  
 PTVLPGHSMKLILUSW"

gene  
 CDS

FEATURES  
 source

COMMENT







4262	A	CAGAATATGTAGTGTCTCCAGTGTCTACGAACACCATGAGAGACACACTCTTTAGA	4321
4322	G	AAGAAGAGAAAACAGGTCCTTGATATCCCAACTGGCAATTGACTTTTCTGATATTA	4381
4322	G	AAGACAG-AAACAGGTCCTTGATATCCCAACTGGCAATTGACTTT--CTGATATAC	4378
4382	A	ACTCTTTTACTGTGCACATGGATGTGCTCTCGAGCCACCATCACTGGCTACAGGAT	4441
4379	A	ACTCTTTTACTGTGCACATGGATGTGCTCTCGAGCCACCATCACTGGCTACAGGAT	4438
4442	C	ATCATCTCCGAGCACTTCAGTGGGAGACCTTCGAGAAGATCGGGTGC	4501
4439	C	ATCATCTCCGAGCACTTCAGTGGGAGACCTTCGAGAAGATCGGGTGC	4498
4502	T	CCATCACCGCTCACCACCTCACTCAGGACAGAGTATGTGGTCA	4561
4499	T	CCATCACCGCTCACCACCTCACTCAGGACAGAGTATGTGGTCA	4558
4562	A	ATGGCAGAGAGAAAGTCCCTTATTTGATTGGCCAACAATCAACAGTTTC	4621
4559	A	ATGGCAGAGAGAAAGTCCCTTATTTGATTGGCCAACAATCAACAGTTTC	4618
4622	A	GGGACCTCGAAGTTGTTGTCGGA	4681
4619	A	GGGACCTCGAAGTTGTTGTCGGA	4678
4682	G	CTGTCA	4741
4679	G	CTGTCA	4738
4742	C	AGGAGTTCACTGTGCTGGGACAA	4801
4739	C	AGGAGTTCACTGTGCTGGGACAA	4798
4802	G	TGATTTACCAATCACTGTGATGCTGTCACTGGCGGTGGAGACAGCCCGC	4861
4799	G	TGATTTACCAATCACTGTGATGCTGTCACTGGCGGTGGAGACAGCCCGC	4858
4862	A	AGCCAAATTTCCATTAATTTACCGAA	4921
4859	A	AGCCAAATTTCCATTAATTTACCGAA	4918
4922	G	ATGTTTCAGGACAAACAGCATTA	4981
4919	G	ATGTTTCAGGACAAACAGCATTA	4978
4982	T	ACAGAGTAACCACTCCCAAAATTTGACAAACCAATCCAGATGCAAGT	5041
4979	T	ACAGAGTAACCACTCCCAAAATTTGACAAACCAATCCAGATGCAAGT	5038
5042	C	CAGATCAAAAGAAATGACTATTGAAGGCTTCAGCCCA	5101
5039	C	CAGATCAAAAGAAATGACTATTGAAGGCTTCAGCCCA	5098
5102	G	TCTATGCTCAGAAATCCAGCGGAGAGTCAGCCTTGTT	5161
5099	G	TCTATGCTCAGAAATCCAGCGGAGAGTCAGCCTTGTT	5158
5162	A	TTGATCGCCCTAAAGACCTGGCAATTCACCTGATGTGGATTCGATTC	5221
5159	A	TTGATCGCCCTAAAGACCTGGCAATTCACCTGATGTGGATTCGATTC	5218
5222	T	GGGAAAGCCACAGGGGCAAGTTTCAGGTACAGGGTGACCTT	5281
5219	T	GGGAAAGCCACAGGGGCAAGTTTCAGGTACAGGGTGACCTT	5278
5282	G	GAATCCATCAGCTATTCCCTGCACCTGATGTGGAAGACACTG	5341
5279	G	GAATCCATCAGCTATTCCCTGCACCTGATGTGGAAGACACTG	5338
5342	C	TACAGCCGGTTCAGATACAGTCAGTGTGGTTGCCCTTGACCATGAT	5401
5339	C	TACAGCCGGTTCAGATACAGTCAGTGTGGTTGCCCTTGACCATGAT	5398

Qy	5402	CAGCCCTGATTGGAAACCCAGTCCACAGCTATTCCTCTGCAACCACTGACCTGAAGTTCACT	5461
Db	5399	CAGCCCTGATTGGAAACCCAGTCCACAGCTATTCCTCTGCAACCACTGACCTGAAGTTCACT	5458
Qy	5462	CAGGTCAACCCCAACAGCCTGAGCGCCAGTGGACACCAACCCAAATGTTCCAGTCTCACTGGA	5521
Db	5459	CAGGTCAACCCCAACAGCCTGAGCGCCAGTGGACACCAACCCAAATGTTCCAGTCTCACTGGA	5518
Qy	5522	TATCGAGTGGGGTGACCCCAAGGAAGAAGACCGGACCAATGAAGAANAATCAACCTTGCT	5581
Db	5519	TATCGAGTGGGGTGACCCCAAGGAAGAAGACCGGACCAATGAAGAANAATCAACCTTGCT	5578
Qy	5582	CCTGACAGCTCATCCCGTGGTGTGATCAGGACTTATGGTGGCCACCAAAATGAAGTGAAGT	5641
Db	5579	CCTGACAGCTCATCCCGTGGTGTGATCAGGACTTATGGTGGCCACCAAAATGAAGTGAAGT	5638
Qy	5642	GTCTATGCTCTTAAGACACTTTGACAAGCAGACCAAGCTCAGGAGTTGTCACCACTCTG	5701
Db	5639	GTCTATGCTCTTAAGACACTTTGACAAGCAGACCAAGCTCAGGAGTTGTCACCACTCTG	5698
Qy	5702	GAGAAATGTCAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACC	5761
Db	5699	GAGAAATGTCAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACC	5758
Qy	5762	ATTAGCTGAGAAACCAAGACTGAGACGATCACTGGCCTCCAGTTGATGCGCTTCAGCC	5821
Db	5759	ATTAGCTGAGAAACCAAGACTGAGACGATCACTGGCCTCCAGTTGATGCGCTTCAGCC	5818
Qy	5822	AATGGCCAGACTCCAATCCAGAGAACCATCAAGCCAGATGTGAGAAGCTACACCATCA	5881
Db	5819	AATGGCCAGACTCCAATCCAGAGAACCATCAAGCCAGATGTGAGAAGCTACACCATCA	5878
Qy	5882	GGCTTACACACCGCACTGACTACAGACTACCTGTACACCTTGAATGACAAATGCTCGG	5941
Db	5879	GGTTTACACACCGCACTGACTACAGACTACCTGTACACCTTGAATGACAAATGCTCGG	5938
Qy	5942	AGCTCCCTGTGGTCTACGACGCTCCACTGCCATTGATGCAACCATCCAACCTGCGTTTC	6001
Db	5939	AGCTCCCTGTGGTCTACGACGCTCCACTGCCATTGATGCAACCATCCAACCTGCGTTTC	5998
Qy	6002	CTGGCCACACACCCAAATTCCTTGCTGGTATCATGTCAGCGCCACAGTCCGCAAGGATTACC	6061
Db	5999	CTGGCCACACACCCAAATTCCTTGCTGGTATCATGTCAGCGCCACAGTCCGCAAGGATTACC	6058
Qy	6062	GGCTACATCATCAAGTATGAGAAGCCTGGGTCTCTCCCCAGAGAAAGTGTCCTCGGCC	6121
Db	6059	GGCTACATCATCAAGTATGAGAAGCCTGGGTCTCTCCCCAGAGAAAGTGTCCTCGGCC	6118
Qy	6122	CGCCCTGGTGTACAGAGCTACTATTACTGCGCTTGGACCGGACCGCAATATACAAT	6181
Db	6119	CGCCCTGGTGTACAGAGCTACTATTACTGCGCTTGGACCGGACCGCAATATACAAT	6178
Qy	6182	TATGTCAATTCGCCCTGAAGAAATTAATCAGAGAGCGAGCCCTGATTGGAAGGAAAAAGACA	6241
Db	6179	TATGTCAATTCGCCCTGAAGAAATTAATCAGAGAGCGAGCCCTGATTGGAAGGAAAAAG---	6235
Qy	6242	GACGAGCTTCCCAACTGGTTAACCTTCCACACCCCAATCTTTCATGGACACAGATCTTG	6301
Db	6236	-----	6235
Qy	6302	GATGTTCTTCCAGTTCAAAAGACCCCTTTGTCACCCACCTTGGGTATGACACTGGA	6361
Db	6236	-----ACAGTTCAAAAGACCCCTTTGTCACCCACCTTGGGTATGACACTGGA	6283
Qy	6362	AATGGTATTACAGCTTCTCGGCACTTCTGTCAGCAACCCAGTGTGGGCAACAAATGATC	6421
Db	6284	AATGGTATTACAGCTTCTCGGCACTTCTGTCAGCAACCCAGTGTGGGCAACAAATGATC	6343
Qy	6422	TTTGAGGAACATGGTTTTAGCGGACCAACCGCCCAACCGGCCACCCCAATAGGCAT	6481
Db	6344	TTTGAGGAACATGGTTTTAGCGGACCAACCGCCCAACCGGCCACCCCAATAGGCAT	6403



Query Match	95.9%	Score 7479.2	DB 9	Length 8320
Best Local Similarity	96.4%	Pred. No. 0		
Matches 7781	Conservative 0	Mismatches 13	Indels 274	Gaps 2
iy	2	GGCTCAGACGGCTGTGCTCCACAGGGGGAGAGGGAAACCCACAGGCCGAGCGGGAGG 61		
ib	1	GGCCCGCGCGGCTGTGCTGTCAAGGGGNGAGAGGGAAACCCAGGCGCGAGCGGGAGG 60		
iy	62	AGGGGACGTGAGGCCAACTTCTGGTCCCTGTGCAATCCCTTCTGTCCCTCCACCCGTC 121		
ib	61	AGGGGACGTGAGGCCAACTTCTCTGGTCTCTGTGCATCCCTTCTGTCCCTCCACCCGTC 120		
iy	122	CCCTTCCCCACCCCTCTGGCCCCCACCTTCTTGAGGCGACAAACCCCGGGAGGCATTAGA 181		
ib	121	CCCTTCCCCACCCCTCTGGCCCCCACCTTCTTGAGGCGCAACACCCCGGGAGGCATTAGA 180		
iy	182	AGGGAATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGGTG 241		
ib	181	AGGGAATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGGTG 240		
iy	242	CGGGCGTCTCTCCCCACCCGCTCAACATGCTTAGGGGTCGGGGCCCGGGCTGTGCTGTG 301		
ib	241	CGGGCGTCTCTCCCCACCCGCTCAACATGCTTAGGGGTCGGGGCCCGGGCTGTGCTGTG 300		
iy	302	CTGCGCGTCTGTGCTCTGGGGAAGCGGTCCTCCACCGGAGCCTTCGAAGAGCAAGAGG 361		
ib	301	CTGCGCGTCTGTGCTCTGGGGAAGCGGTCCTCCACCGGAGCCTTCGAAGAGCAAGAGG 360		
iy	362	CAGGCTCAGCAATATGTTTACGCCCCAGTCCCGGTGCTGTGATCAAGCAAGGCCCGGT 421		
ib	361	CAGGCTCAGCAATATGTTTACGCCCCAGTCCCGGTGCTGTGATCAAGCAAGGCCCGGT 420		
iy	422	TGTTATGCAATAGAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTTACTAGGC 481		
ib	421	TGTTATGCAATAGAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTTACTAGGC 480		





3721 ACCATCCAGTCTGAGAGTGGACAGGAAAGAGATGCGCCAAATGTTAAACAAGTGTG 3780  
3782 ACACCAATGCTCCACCAACAAACTTGCATCTCGAGGCAAAACCTGACACTGGATGCTC 3841  
3781 ACACCAATGCTCCACCAACAAACTTGCATCTCGAGGCAAAACCTGACACTGGATGCTC 3840  
3842 ACAGTCTCTGGAGAGAGACACACCCGACATTAATGATGTTATAGAAATTAACCAACCC 3901  
3841 ACAGTCTCTGGAGAGAGACACACCCGACATTAATGATGTTATAGAAATTAACCAACCC 3900  
3902 CCTACAAACGGCCAGCAGGAAATCTTT - GGAAAGAGTGGTCCATGCTGATCAGAGTTC 3960  
3901 CCTACAAACGGCCAGCAGGAAATCTTT - GGAAAGAGTGGTCCATGCTGATCAGAGTTC 3960  
3961 CTGCACTTTTGATAA CTTGAGTCCCGCTCGGAGTACAATGTCAGTGTATACAGTCAA 4020  
3961 CTGCACTTTTGATAA CTTGAGTCCCGCTCGGAGTACAATGTCAGTGTATACAGTCAA 4020  
4021 GGATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGAGTCCCGAATCMAC 4063  
4021 GGATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGAGTCCCGAATCMAC 4080  
4064 - - - - - 4063  
4081 TGACCTAAGCTTTGTTGATATAACCGATTCAAGCATCGGCTGAGGTGAGCCGCTAAA 4140  
4064 - - - - - 4063  
4141 CTCTTCCACCATTAATGGGTACCGCATACAGTAGTGGCGAGGAAAGTATCCCTAT 4200  
4064 - - - - - 4063  
4201 TTTTGAAGATTTTGTGACTCTCTCAGTAGTACTACACAGTACACAGGCTGGAGCGGG 4260  
4064 - - - - - 4063  
4261 CATGACTATGATATCAGCGTTATCACTCTCAATTAATGGCGGAGAGTGCCTTACTAC 4320  
4064 - - - - - 4107  
4321 ACTGACACACAAACGGCTGTCTCTCTCCACTGACCTGCGATTCACCAACATTTGGTCC 4380  
4108 AGACACATGCTGTGACCTGGGCTGACCCCATCATGATGATTAACCACTTCCTGTGT 4167  
4381 AGACACATGCTGTGACCTGGGCTGACCCCATCATGATGATTAACCACTTCCTGTGT 4440  
4168 GCCTTACTACCTGTGAAATACAGGAAAGTGTGAGAGTGTGCAATTTCTCTTTCAGA 4227  
4441 GCCTTACTACCTGTGAAATACAGGAAAGTGTGAGAGTGTGCAATTTCTCTTTCAGA 4500  
4228 CAATGCAAGTGTCTTAACAAATCTCTGCTGTGATCAGAAATATGTAAGTGTCTCCAG 4287  
4501 CAATGCAAGTGTCTTAACAAATCTCTGCTGTGATCAGAAATATGTAAGTGTCTCCAG 4560  
4288 TGTCTAGAACACATCAGAGCACCTCTTAGAGAGACAGAAACAGGCTTGTATTC 4347  
4561 TGTCTAGAACACATCAGAGCACCTCTTAGAGAGACAGAAACAGGCTTGTATTC 4620  
4348 CCCAATGCAATTTGATTTCTGATATTAATGTCGCAACTCTTTTACTGTGCACTGGATTGC 4407  
4621 CCCAATGCAATTTGATTTCTGATATTAATGTCGCAACTCTTTTACTGTGCACTGGATTGC 4680  
4408 TCCTGAGCCACCATCTGCTGTACAGGATCCGCTATCCGAGCATCTCAGTGGAG 4467  
4681 TCCTGAGCCACCATCTGCTGTACAGGATCCGCTATCCGAGCATCTCAGTGGAG 4740  
4468 ACCTCGAAGAGATCGGGTGCCTTCTGATATTAATGTCGCAACTCTTTTACTGTGCACTGGATTGC 4527  
4741 ACCTCGAAGAGATCGGGTGCCTTCTGATATTAATGTCGCAACTCTTTTACTGTGCACTGGATTGC 4800  
4528 AGGACAGAGTATGCTGACAGATCTGTTCTTCTTATGAGAGAGAAAGTCCCTTAT 4587  
4801 AGGACAGAGTATGCTGACAGATCTGTTCTTCTTATGAGAGAGAAAGTCCCTTAT 4860

4588 GATTGGCCAAACAATCAACAGATTTCTGATGTTCCGAGGGAACCTGGAAAGTGTGCTGGAC 4647  
4861 GATTGGCCAAACAATCAACAGATTTCTGATGTTCCGAGGGAACCTGGAAAGTGTGCTGGAC 4920  
4648 CCCACACAGCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAATATACAGAT 4707  
4921 CCCACACAGCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAATATACAGAT 4980  
4708 CACTTAGGGAACACAGAGGAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAA 4767  
4981 CACTTAGGGAACACAGAGGAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAA 5040  
4768 GTCTACAGCTACCATCAGCGGCTTAAACCTGAGTGTATATACCATCACTGTGTATGC 4827  
5041 GTCTACAGCTACCATCAGCGGCTTAAACCTGAGTGTATATACCATCACTGTGTATGC 5100  
4828 TGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAATTTCCATTAATACCGAAC 4887  
5101 TGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAATTTCCATTAATACCGAAC 5160  
4888 AGAAATGACAAACCATCCAGATGCAAGTGAACGATGTTTCAGGACAAACAGATAGTGT 4947  
5161 AGAAATGACAAACCATCCAGATGCAAGTGAACGATGTTTCAGGACAAACAGATAGTGT 5220  
4948 CAAGTGGCTGCTTCAAGTTCCCTGTTTACTGTTTACAGAGTAAACCACTCCCAAAA 5007  
5221 CAAGTGGCTGCTTCAAGTTCCCTGTTTACTGTTTACAGAGTAAACCACTCCCAAAA 5280  
5008 TGGACAGGACCAACAAAACTAAAACTGCAAGTTCAGATCAACACAGAAATGACTATTGA 5067  
5281 TGGACAGGACCAACAAAACTAAAACTGCAAGTTCAGATCAACACAGAAATGACTATTGA 5340  
5068 AGGCTTCAGCCACAGTGGAGTATGTTGTTAGTGTCTATGCTCAGAAATCCAAAGCGAGA 5127  
5341 AGGCTTCAGCCACAGTGGAGTATGTTGTTAGTGTCTATGCTCAGAAATCCAAAGCGAGA 5400  
5128 GAGTCAAGCTCTGTTTCAAGTTCAGATGCAAGTAAACCAATGATGCGCCCTAAAGCACTGGCATT 5187  
5401 GAGTCAAGCTCTGTTTCAAGTTCAGATGCAAGTAAACCAATGATGCGCCCTAAAGCACTGGCATT 5460  
5188 CACTGATGCTGATGCTGATTTCCATCBAATTTGTTGGGAAGCCACAGGGGCAAGTTTC 5247  
5461 CACTGATGCTGATGCTGATTTCCATCBAATTTGTTGGGAAGCCACAGGGGCAAGTTTC 5520  
5248 CAGGTACAGGCTGACCTTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGCAAC 5307  
5521 CAGGTACAGGCTGACCTTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGCAAC 5580  
5308 TGATGTTGAAGAGACACATGCAAGCTGCAAGGCTCAGACCCGGTTCAGATACAGT 5367  
5581 TGATGTTGAAGAGACACATGCAAGCTGCAAGGCTCAGACCCGGTTCAGATACAGT 5640  
5368 CAGTGTGCTGCTTGCACATGATATGAGAGAGCCAGCCCTGATGGAACCCAGTCCAC 5427  
5641 CAGTGTGCTGCTTGCACATGATATGAGAGAGCCAGCCCTGATGGAACCCAGTCCAC 5700  
5428 AGCTATTCTGCAACCACTGACCTGAAAGTTCATCAGGTCAACCCACAGCTGAGGCG 5487  
5701 AGCTATTCTGCAACCACTGACCTGAAAGTTCATCAGGTCAACCCACAGCTGAGGCG 5760  
5488 CAGTGTGACACCCCAATGTTTCACTGATATCGAGTGGGGTGAACCCCAAGGA 5547  
5761 CAGTGTGACACCCCAATGTTTCACTGATATCGAGTGGGGTGAACCCCAAGGA 5820  
5548 GAAGACCGGACCAATGAAGAAATCAACCTTGTCTTGAAGCTCATCCGTTGTTATC 5607  
5821 GAAGACCGGACCAATGAAGAAATCAACCTTGTCTTGAAGCTCATCCGTTGTTATC 5880  
5608 AGGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAAGGACACTTTGAC 5667  
5881 AGGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAAGGACACTTTGAC 5940

QY 5668 AAGCAGACCACTCAGGAGTTGTACCACTCTGGAGAAATGTGACGCCAACAAGAGGGC 5727  
DB 5941 AAGCAGACCACTCAGGAGTTGTACCACTCTGGAGAAATGTGACGCCAACAAGAGGGC 6000  
QY 5728 TCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGAC 5787  
DB 6001 TCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGAC 6060  
QY 5788 GATCCTGGCTCCCAAGTTGATGCGTTTCAGGCAATGGCCAGACTCCAAATCCAGAGAAC 5847  
DB 6061 GATCCTGGCTCCCAAGTTGATGCGTTTCAGGCAATGGCCAGACTCCAAATCCAGAGAAC 6120  
QY 5848 CATCAAGCCAGATGTCAGAGCTACACCATCAGCTGGCTTACCAACCAAGGCACTGACTACAA 5907  
DB 6121 CATCAAGCCAGATGTCAGAGCTACACCATCAGCTGGCTTACCAACCAAGGCACTGACTACAA 6180  
QY 5908 GATCCTGGCTCCCAAGTTGATGCGTTTCAGGCAATGGCCAGACTCCAAATCCAGAGAAC 5967  
DB 6181 GATCCTGGCTCCCAAGTTGATGCGTTTCAGGCAATGGCCAGACTCCAAATCCAGAGAAC 6240  
QY 5968 CACTGCGATTGATGACCATCCAACTGCTTCTTGGCCACCAACCCCAATTCCTTGGCT 6027  
DB 6241 CACTGCGATTGATGACCATCCAACTGCTTCTTGGCCACCAACCCCAATTCCTTGGCT 6300  
QY 6028 GGTATCATGGCAGCGCCGACCGTGGCAGGATTACCGGCTTACATCATCAAGTATGAGAGCC 6087  
DB 6301 GGTATCATGGCAGCGCCGACCGTGGCAGGATTACCGGCTTACATCATCAAGTATGAGAGCC 6360  
QY 6088 TGGGTCTCTCCAGAGAAAGTGTCTTCTTGGCCGCGGCTTGTGTACAGAGGCTACTAT 6147  
DB 6361 TGGGTCTCTCCAGAGAAAGTGTCTTCTTGGCCGCGGCTTGTGTACAGAGGCTACTAT 6420  
QY 6148 TACTGGCTGGAAACCGGAAACCGAATATACAAATTTATGTCTATTCCTTGAAGAAATATCA 6207  
DB 6421 TACTGGCTGGAAACCGGAAACCGAATATACAAATTTATGTCTATTCCTTGAAGAAATATCA 6480  
QY 6208 GAAGAGGAGCCCTGTGATGAGAGAAAGAGACAGAGCTTCCCAACCTGAGTAACTTCA 6267  
DB 6481 GAAGAGGAGCCCTGTGATGAGAGAAAGAGACAGAGCTTCCCAACCTGAGTAACTTCA 6540  
QY 6268 TCCACACCCCAATCTTATGAGACAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGAC 6327  
DB 6541 TCCACACCCCAATCTTATGAGACAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGAC 6600  
QY 6328 CCCTTTCTGTCACCCACCTCGGTATGACACTGGAATGGTATTCAGCTTCTCGGACTTTC 6387  
DB 6601 CCCTTTCTGTCACCCACCTCGGTATGACACTGGAATGGTATTCAGCTTCTCGGACTTTC 6660  
QY 6388 TGGTCAGCAACCCAGTGTGGCAACCAATGATCTTTCAGAGAACATGGTTTTCAGCGGAC 6447  
DB 6661 TGGTCAGCAACCCAGTGTGGCAACCAATGATCTTTCAGAGAACATGGTTTTCAGCGGAC 6720  
QY 6448 CACACCGCCCAACACCGGCCACCCCAATAGGCATAGGCCAAGACCAATCCCGCCGAATGT 6507  
DB 6721 CACACCGCCCAACACCGGCCACCCCAATAGGCATAGGCCAAGACCAATCCCGCCGAATGT 6780  
QY 6508 AGGACAGAGCTCTCTCAGACCAACCATCTCATGGGCCCAATTCAGGACACTTCTGA 6567  
DB 6781 AGGACAGAGCTCTCTCAGACCAACCATCTCATGGGCCCAATTCAGGACACTTCTGA 6840  
QY 6568 GTACATCATTTTCATGTCATCTGTTGGCACTGATGAAGAACCCCTTACAGTTTCAGGGTTC 6627  
DB 6841 GTACATCATTTTCATGTCATCTGTTGGCACTGATGAAGAACCCCTTACAGTTTCAGGGTTC 6900  
QY 6628 TGGAACTTCTACAGTGGCACTCTGACAGGCTCTACAGAGTGGCCACTACACATCAT 6687  
DB 6901 TGGAACTTCTACAGTGGCACTCTGACAGGCTCTACAGAGTGGCCACTACACATCAT 6960  
QY 6688 AGTGGAGGCACTGAAGAGACAGCAGAGGATAGGTTTCGGAGAGGTTGTTCAGGTTGG 6747  
DB 6961 AGTGGAGGCACTGAAGAGACAGCAGAGGATAGGTTTCGGAGAGGTTGTTCAGGTTGG 7020  
QY 6748 CAACTCTGTCAAGAGAGGCTTGAACCAACCTAGGATGATCTGCTGCTTGGACCCCTACAC 6807

RESULT 8

DB 7021 CAACTCTGTCAACGAGGCTTGAACCAACCTACGATGACTCGTGTGTTGACCCCTACAC 7080  
QY 6808 AGTGCCCACTATATGCGGTTGGAGATGAGTGGGAAACGAATGTCTGAATCAGGCTTTAAACT 6867  
DB 7081 AGTTCCTCCATATATGCGGTTGGAGATGAGTGGGAAACGAATGTCTGAATCAGGCTTTAAACT 7140  
QY 6868 GTTGTGCCAGTGTCTTGGCTTTGGAGTGGTCAATTCAGATGTGATTCATCTAGATGTTG 6927  
DB 7141 GTTGTGCCAGTGTCTTGGCTTTGGAGTGGTCAATTCAGATGTGATTCATCTAGATGTTG 7200  
QY 6928 CCATGACCAATGTGTGAACCTACAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGG 6987  
DB 7201 CCATGACCAATGTGTGAACCTACAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGG 7260  
QY 6988 CAGATGATGAGCTGCAATGTCTTGGGAAACGGAAGAGAGATTCAGTGTGACCTCA 7047  
DB 7261 CAGATGATGAGCTGCAATGTCTTGGGAAACGGAAGAGAGATTCAGTGTGACCTCA 7320  
QY 7048 TGAGGCAACGCTTACGATGATGGGAAACACATACCACCTAGGAGAACAGTGGCAGAGGA 7107  
DB 7321 TGAGGCAACGCTTATGATGATGGGAAACACATACCACCTAGGAGAACAGTGGCAGAGGA 7380  
QY 7108 ATATCTCGTGTCCATTTGCTCTGCAATGCTTTGGAGCCAGCGGCTGGCGCTGTGA 7167  
DB 7381 ATATCTCGTGTCCATTTGCTCTGCAATGCTTTGGAGCCAGCGGCTGGCGCTGTGA 7440  
QY 7168 CAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGCCTACTCTGCGCCAGTCTCAAA 7227  
DB 7441 CAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGCCTACTCTGCGCCAGTCTCAAA 7500  
QY 7228 CAGATATCTCAGAGATACCATCAGAGAACAAACATAATGTTAATTTGCCAAATGAGTG 7287  
DB 7501 CAGATATCTCAGAGATACCATCAGAGAACAAACATAATGTTAATTTGCCAAATGAGTG 7560  
QY 7288 CTTTCATGCTTTTGTAGATGATCAGGCTGACAGAGAAATTTCCCGAGAGTAAATCATCTTCC 7347  
DB 7561 CTTTCATGCTTTTGTAGATGATCAGGCTGACAGAGAAATTTCCCGAGAGTAAATCATCTTCC 7620  
QY 7348 AATCCAGAGAAACAGCATGTCTCTGCAAGATTCATCTTAAATCTGAGTGATGTTAGC 7407  
DB 7621 AATCCAGAGAAACAGCATGTCTCTGCAAGATTCATCTTAAATCTGAGTGATGTTAGC 7680  
QY 7408 AGACCCAGCTTGTAGCTTCTTCTTCTTAAAGCCCTTCTGCTCTGGAGGAGTCTCCA 7467  
DB 7681 AGACCCAGCTTGTAGCTTCTTCTTCTTAAAGCCCTTCTGCTCTGGAGGAGTCTCCA 7740  
QY 7468 GCTTCAGCTCAACTCAGCTTCTCCAGCATCACCTCTGGAGTTCCTGAGGGTCTTCT 7527  
DB 7741 GCTTCAGCTCAACTCAGCTTCTCCAGCATCACCTCTGGAGTTCCTGAGGGTCTTCT 7800  
QY 7528 CATTAATGAGGGGTGACATTCCTGCTTCTGCTTGAAGTATTCATACCGCTCAGTATT 7587  
DB 7801 CATTAATGAGGGGTGACATTCCTGCTTCTGCTTGAAGTATTCATACCGCTCAGTATT 7860  
QY 7588 TTAATTAAGTGTATTCAGATTTGGTGGATCAATAGGAAAGCATATGAGCCCAACC 7647  
DB 7861 TTAATTAAGTGTATTCAGATTTGGTGGATCAATAGGAAAGCATATGAGCCCAACC 7920  
QY 7648 AGATGCAAAATGTTTGAATGATGACCAAAAATTTAAGTAGGAAAGTCAACCAACA 7707  
DB 7921 AAGATGCAAAATGTTTGAATGATGATGACCAAAAATTTAAGTAGGAAAGTCAACCAACA 7980  
QY 7708 CTTCTGCTTCTCACTTAAGTGTCTGGCCCGCAATATCTGTAGGAAACAGCATGATCTTTGA 7767  
DB 7981 CTTCTGCTTCTCACTTAAGTGTCTGGCCCGCAATATCTGTAGGAAACAGCATGATCTTTGA 8040  
QY 7768 CTGTGATATTTTAAATATCCACAGTACT 7795  
DB 8041 CTGTGATATTTTAAATATCCACAGTACT 8068









6181 GACACCATCTCATGGGCCCCATCCAGGACACCTCTGAGTACATCATTTTCATGTCATCC 6240  
 QY  
 6589 TGTGGGCACTGATGAAGAACCCCTTACAGTTCAGGGTTCCTGGAACTTCTACCACTGCCAC 6648  
 Db  
 6241 TGTGGGCACTGATGAAGAACCCCTTACAGTTCAGGGTTCCTGGAACTTCTACCACTGCCAC 6300  
 QY  
 6649 TCTGACAGCCCTCACAGAGGTCGCCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6708  
 Db  
 6301 TCTGACAGCCCTCACAGAGGTCGCCACCTACAAATCATAGTGGAGGCACTGAAAGACCA 6360  
 QY  
 6709 GCAGAGGATAGGTTCCGGGAAGAGGTTGTTACCGTGGGGAACCTCTGTCAACGAGGCTT 6768  
 Db  
 6361 GCAGAGGATAGGTTCCGGGAAGAGGTTGTTACCGTGGGGAACCTCTGTCAACGAGGCTT 6420  
 QY  
 6769 GAAACCAACTACCGATGACTCGTGTCTTGAACCCCTACACAGTGTCCCATATATGCGGTGG 6828  
 Db  
 6421 GAAACCAACTACCGATGACTCGTGTCTTGAACCCCTACACAGTGTCCCATATATGCGGTGG 6480  
 QY  
 6829 AGATGAGTGGGAACCAATGTCTGAATCAAGGCTTTAACTGTTGGCAGTCTTAGGCTT 6888  
 Db  
 6481 AGATGAGTGGGAACCAATGTCTGAATCAAGGCTTTAACTGTTGGCAGTCTTAGGCTT 6540  
 QY  
 6889 TGGAGTGTCTCATTTTCAGATGTGATTCATCTAGATGGTGCATGACATGTGTGTAACCTA 6948  
 Db  
 6541 TGGAGTGTCTCATTTTCAGATGTGATTCATCTAGATGGTGCATGACATGTGTGTAACCTA 6600  
 QY  
 6949 CAAGATTGGAGAGAAAGTGGGACCGTTCAGGGAAGAAATGGCCAGATGATGAGCTGCAATG 7008  
 Db  
 6601 CAAGATTGGAGAGAAAGTGGGACCGTTCAGGGAAGAAATGGCCAGATGATGAGCTGCAATG 6660  
 QY  
 7009 TCTTGGGAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCAACGTGTACGATGA 7068  
 Db  
 6661 TCTTGGGAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCAACGTGTACGATGA 6720  
 QY  
 7069 TGGGAAGACATACCACTAGGAGAACAGTGGGAGAAATATCTCGGTGCCATTTGCTC 7128  
 Db  
 6721 TGGGAAGACATACCACTAGGAGAACAGTGGGAGAAATATCTCGGTGCCATTTGCTC 6780  
 QY  
 7129 CTGACATGCTTTGGAGCCAGCGGGTGGCGTGTGACAACTGCCGACAGCTGGGG 7188  
 Db  
 6781 CTGACATGCTTTGGAGCCAGCGGGTGGCGTGTGACAACTGCCGACAGCTGGGG 6840  
 QY  
 7189 TGAACCCAGTCCCGAAGGCACTACTGCGGAGTCTTACAAACAGATATCTTCAGAGATACCA 7248  
 Db  
 6841 TGAACCCAGTCCCGAAGGCACTACTGCGGAGTCTTACAAACAGATATCTTCAGAGATACCA 6900  
 QY  
 7249 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTCATGCCCTTTAGATGTACA 7308  
 Db  
 6901 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTCATGCCCTTTAGATGTACA 6960  
 QY  
 7309 GGTGACAGAGAGATTCGAGAGTAAATCATCTTCCAAATCCAGAGGAAACAGCATGT 7368  
 Db  
 6961 GGTGACAGAGAGATTCGAGAGTAAATCATCTTCCAAATCCAGAGGAAACAGCATGT 7020  
 QY  
 7369 CTCTGCGCAAGATCCATCTAACTGGAGTGTATGAGAGCCAGCTTAGAGTTCTTC 7428  
 Db  
 7021 CTCTGCGCAAGATCCATCTAACTGGAGTGTATGAGAGCCAGCTTAGAGTTCTTC 7080  
 QY  
 7429 TTTCTTTCTTAAGCCCTTGTCTGAGGAGTGTCTCCAGCTTCACTCACTCAGACT 7488  
 Db  
 7081 TTTCTTTCTTAAGCCCTTGTCTGAGGAGTGTCTCCAGCTTCACTCACTCAGACT 7140  
 QY  
 7489 TCTCCAGCATCACCTCGGAGTTCTCGAGGTTTCTCATAAATGAGGGCTGCACATT 7548  
 Db  
 7141 TCTCCAGCATCACCTCGGAGTTTCTCGAGGTTTCTCATAAATGAGGGCTGCACATT 7200  
 QY  
 7549 GCTGTGTTCTGCTTCAAGTATTCATACCGCTCAGTATTTAAATGAGTGTATTTCTAAGA 7608  
 Db  
 7201 GCTGTGTTCTGCTTCAAGTATTCATACCGCTCAGTATTTAAATGAGTGTATTTCTAAGA 7260  
 QY  
 7609 TTTGGTTGGGATCAATAGGAAGCATATGCGAGCCAAACAGATCCAAATGTTTGAAT 7668  
 Db  
 7261 TTTGGTTGGGATCAATAGGAAGCATATGCGAGCCAAACAGATCCAAATGTTTGAAT 7320

QY 7669 GATATGACCAATATTTTAAGTAGGAAAGTCAACCAACACCTTCTGCTTTCACTTAAGTGT 7728  
 Db 7321 GATATGACCAATATTTTAAGTAGGAAAGTCAACCAACACCTTCTGCTTTCACTTAAGTGT 7380  
 QY 7729 CTGGCCCGCAATACCTGTAGGAAACAGCATGATCTTTGTTACTGTGATATTTTAAATATCA 7788  
 Db 7381 CTGGCCCGCAATACCTGTAGGAAACAGCATGATCTTTGTTACTGTGATATTTTAAATATCA 7440  
 QY 7789 CAGTACT 7795  
 Db 7441 CAGTACT 7447

RESULT 9  
 AR380744  
 LOCUS AR380744 7680 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 1289 from patent US 6607879.  
 ACCESSION AR380744  
 VERSION AR380744.1 GI:40088378  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 7680)  
 AUTHORS Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.  
 TITLE Compositions for the detection of blood cell and immunological response gene expression  
 JOURNAL Patent: US 6607879-A 1289 19-AUG-2003;  
 FEATURES  
 source 1..7680  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 95.2%; Score 7423; DB 6; Length 7680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 349 GAAGACCAAGAGCGAGCTCAGCAATGTTTTCAGCCCAAGTCCCGGTGGCTGTCACTCA 408  
 Db 1 GAAGACCAAGAGCGAGCTCAGCAATGTTTTCAGCCCAAGTCCCGGTGGCTGTCACTCA 60  
 QY 409 AAGCAAGCCCGTGTGTTATGACAAATGGAATAAATCACTATCAGATAAATCAACAGTGGAGCG 468  
 Db 61 AAGCAAGCCCGTGTGTTATGACAAATGGAATAAATCACTATCAGATAAATCAACAGTGGAGCG 120  
 QY 469 GACCTACCTAGGCAATGCGCTTGGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTG 528  
 Db 121 GACCTACCTAGGCAATGCGCTTGGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTG 180  
 QY 529 CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTTGGGAAACACTTACCG 588  
 Db 181 CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTTGGGAAACACTTACCG 240  
 QY 589 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACCTGATCCG 648  
 Db 241 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACCTGATCCG 300  
 QY 649 GGTGGGCGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAAAGGGGCTCAGTC 708  
 Db 301 GGTGGGCGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAAAGGGGCTCAGTC 360  
 QY 709 CTAACAGATTGGTGACACCTGGAGGAGACCAATGAGACTGTGGTGTACATGTTAGAGTG 768  
 Db 361 CTAACAGATTGGTGACACCTGGAGGAGACCAATGAGACTGTGGTGTACATGTTAGAGTG 420  
 QY 769 TGTGTGCTTGTGTAATGGAAGAGAAATGGAGCTTCCAAAGCCCATAGCTGAGAAAGTGT 828  
 Db 421 TGTGTGCTTGTGTAATGGAAGAGAAATGGAGCTTCCAAAGCCCATAGCTGAGAAAGTGT 480  
 QY 829 TGATCATGCTGTGGGACTTCTTATGTTGCGAGAAACGTTGGGAGAACCTTACCAAG 888









4861 CATCAAAATTTGCTTGGGAAGGCCCAAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 4920  
5269 GAGCCCTGAGATGGAATCCATGAGCTATTCCCTGACACCTGATGTTGAAGAAAGACACTGC 5328  
4921 GAGCCCTGAGATGGAATCCATGAGCTATTCCCTGACACCTGATGTTGAAGAAAGACACTGC 4980  
5329 AGAGCTGCAGAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTTGGTTCCTTGCACGA 5388  
4981 AGAGCTGCAGAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTTGGTTCCTTGCACGA 5040  
5389 TGATATGGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTTGCACCAACTGA 5448  
5041 TGATATGGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTTGCACCAACTGA 5100  
5449 CTTGAAGTTTCACTCAGTTCACACCCACAGCCCTGAGCGCCAGTGACACACCCCAATGT 5508  
5101 CTTGAAGTTTCACTCAGTTCACACCCACAGCCCTGAGCGCCAGTGACACACCCCAATGT 5160  
5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGA 5568  
5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGA 5220  
5569 AATCAACCTTGTCTTGCAGAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCA 5628  
5221 AATCAACCTTGTCTTGCAGAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCA 5280  
5629 ATATGAAGTGAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACCACTCAGGGAGT 5688  
5281 ATATGAAGTGAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACCACTCAGGGAGT 5340  
5689 TGTCAACACTCTGGAGATGTGACGCCCAAGAGGGCTCGTGTGACAGATGCTACTGA 5748  
5341 TGTCAACACTCTGGAGATGTGACGCCCAAGAGGGCTCGTGTGACAGATGCTACTGA 5400  
5749 GACCACTCACTTATGCTGGAGACCAAGACTGAGACGATCAGTGGCTCCAGTTGA 5808  
5401 GACCACTCACTTATGCTGGAGACCAAGACTGAGACGATCAGTGGCTCCAGTTGA 5460  
5809 TGCCGTTTCCAGCAATGGCCAGACTCCAATCCAGAGAACCATCAAGCCAGATGTCAAG 5868  
5461 TGCCGTTTCCAGCAATGGCCAGACTCCAATCCAGAGAACCATCAAGCCAGATGTCAAG 5520  
5869 CTACACACTCTGGCTTACACACAGGCTGACTACAGATCTACCTGTACACTTGA 5928  
5521 CTACACACTCTGGCTTACACACAGGCTGACTACAGATCTACCTGTACACTTGA 5580  
5929 TGCAATGCTCGAGCTCCCTGCTGCTACAGGCTCCACTGCCATGATGACCACTC 5988  
5581 TGCAATGCTCGAGCTCCCTGCTGCTACAGGCTCCACTGCCATGATGACCACTC 5640  
5989 CAACCTGCTGCTGCTCCCTGCTGCTACAGGCTCCACTGCCATGATGACCACTC 6048  
5641 CAACCTGCTGCTGCTCCCTGCTGCTACAGGCTCCACTGCCATGATGACCACTC 5700  
6049 TGCCAGATTTACCGGCTACATCAATCAAGTATGAGAGCTTGGTCTCTCCAGAGAGT 6108  
5701 TGCCAGATTTACCGGCTACATCAATCAAGTATGAGAGCTTGGTCTCTCCAGAGAGT 5760  
6109 GGTCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6168  
5761 GGTCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5820  
6169 CGAATATCAATTTATGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6228  
5821 CGAATATCAATTTATGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5880  
6229 AAGGAAAGAGACAGAGCTTCCCACTGGTAACTTCCACACCCCAATCTTCTATGG 6288  
5881 AAGGAAAGAGACAGAGCTTCCCACTGGTAACTTCCACACCCCAATCTTCTATGG 5940  
6289 ACCAGAGATCTTGGATGCTTCCAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 6348  
5941 ACCAGAGATCTTGGATGCTTCCAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 6000

6349 GTATGACACTGGAATGGTATTTCAGCTTCTGGGCACTTCTGGTCAAGAACCCAGTGTGG 6408  
6001 GTATGACACTGGAATGGTATTTCAGCTTCTGGGCACTTCTGGTCAAGAACCCAGTGTGG 6060  
6409 GCACAAATGATCTTTCAGGACACATGGTTCCTGAGGACACACACCGCCACACAGGCGAC 6468  
6061 GCACAAATGATCTTTCAGGACACATGGTTCCTGAGGACACACACCGCCACACAGGCGAC 6120  
6469 CCCCATAGGCTAGGCAAGACCATACCGCCGCAATGTAGGACAAAGAGCTCTCTCTCA 6528  
6121 CCCCATAGGCTAGGCAAGACCATACCGCCGCAATGTAGGACAAAGAGCTCTCTCTCA 6180  
6529 GACAAACATCTCATGAGGCCCCCATTCAGGACACATTCAGAGTACATCATTTTCATCTATCC 6588  
6181 GACAAACATCTCATGAGGCCCCCATTCAGGACACATTCAGAGTACATCATTTTCATCTATCC 6240  
6589 TGTTCGCACTGATCAAGAACCCCTTACAGTTCAGGGTTCCTGGAACTTCTACCACTGGCCAC 6648  
6241 TGTTCGCACTGATGAAGAACCCCTTACAGTTCAGGGTTCCTGGAACTTCTACCACTGGCCAC 6300  
6649 TCTGACAGGCTCTACACAGAGGTGACCTACACATCATATAGTGGAGGACCTGAAGACCA 6708  
6301 TCTGACAGGCTCTACACAGAGGTGACCTACACATCATATAGTGGAGGACCTGAAGACCA 6360  
6709 GCAGAGGCTATAGGTTTCGGGAAGAGGTTTACCGTGGGCAACTCTGTCAACGAGGCTT 6768  
6361 GCAGAGGCTATAGGTTTCGGGAAGAGGTTTACCGTGGGCAACTCTGTCAACGAGGCTT 6420  
6769 GAAACAACTTACAGATGACTGCTGCTTTCAGGCTTACACAGTGTCCCATTTATGCGGTGG 6828  
6421 GAAACAACTTACAGATGACTGCTGCTTTCAGGCTTACACAGTGTCCCATTTATGCGGTGG 6480  
6829 AGATGAGTGGGAACGAAATGCTGAATCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTT 6888  
6481 AGATGAGTGGGAACGAAATGCTGAATCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTT 6540  
6889 TGAAGTGGTTCATTTACAGATGCTGATTCATCTAGATGCTGCTGCTGCTGCTGCTGCTGCT 6948  
6541 TGAAGTGGTTCATTTACAGATGCTGATTCATCTAGATGCTGCTGCTGCTGCTGCTGCTGCT 6600  
6949 CAAGATTTGGAGAGGTTGGGACCGTTCAGGAGAGAAATGGCCAGATGATGAGTGCACATG 7008  
6601 CAAGATTTGGAGAGGTTGGGACCGTTCAGGAGAGAAATGGCCAGATGATGAGTGCACATG 6660  
7009 TCTTTGGGAACCGAAAAGGAGAAATTCAGAGTGTGACCCCTCATGAGGCAACGTTTACGATGA 7068  
6661 TCTTTGGGAACCGAAAAGGAGAAATTCAGAGTGTGACCCCTCATGAGGCAACGTTTACGATGA 6720  
7069 TGGGAAGACATACACAGTAGGAGACAGTGGCAGAGAAATATCTCGTGGCCATTTGCTC 7128  
6721 TGGGAAGACATACACAGTAGGAGAACAGTGGCAGAGAAATATCTCGTGGCCATTTGCTC 6780  
7129 CTGCACATGCTTTTGGAGGCCAGCGGGCTGCGCTGTGACAACTGCGCCAGACCTGGGGG 7188  
6781 CTGCACATGCTTTTGGAGGCCAGCGGGCTGCGCTGTGACAACTGCGCCAGACCTGGGGG 6840  
7189 TGAACCCAGTCCCGAAGGCACTACTGCGCAGTCTTCAACACAGTATTTCTCAGAGATACCA 7248  
6841 TGAACCCAGTCCCGAAGGCACTACTGCGCAGTCTTCAACACAGTATTTCTCAGAGATACCA 6900  
7249 TCAGAGAACAAACACTAATGTTAATTCGCCAATTTGAGTGTCTCATGCTTTAGATGTACA 7308  
6901 TCAGAGAACAAACACTAATGTTAATTCGCCAATTTGAGTGTCTCATGCTTTAGATGTACA 6960  
7309 GGTTCAGAGAGAGATTTCCGAGAGTAAATCATCTTTTCCCAATCCAGAGGAAACAGCATGT 7368  
6961 GGTTCAGAGAGAGATTTCCGAGAGTAAATCATCTTTTCCCAATCCAGAGGAAACAGCATGT 7020  
7369 CTCTCTGCCAGATTCATCTAACTTGAAGTGTAGCAGACCCAGCTTAGAGTTCCTTC 7428  
7021 CTCTCTGCCAGATTCATCTAACTTGAAGTGTAGCAGACCCAGCTTAGAGTTCCTTC 7080

```
QY 7429 TTTCTTTCTTAAGCCCTTTGCTCTGGAGAAAGTTCTCCAGCTTCAGCTCAACTCACAGCT 7488
Db 7081 TTTCTTTCTTAAGCCCTTTGCTCTGGAGAAAGTTCTCCAGCTTCAGCTCAACTCACAGCT 7140
QY 7489 TCTCCAGCANTCACCTGGAGTTTCTCTGAGGTTTCTCATAAATGAGGGCTGCACATT 7548
Db 7141 TCTCCAGCANTCACCTGGAGTTTCTCTGAGGTTTCTCATAAATGAGGGCTGCACATT 7200
QY 7549 GCCTCTTCTGCTTCCGAAGTATTCAATACCGCTCAGTATTTTAAATGAAGTGATTTCTAAGA 7608
Db 7201 GCCTCTTCTGCTTCCGAAGTATTCAATACCGCTCAGTATTTTAAATGAAGTGATTTCTAAGA 7260
QY 7609 TTTGGTTTGGATCAATAGGAAGCATATGAGCAACCAACCATGCAAAATGTTTGAAT 7668
Db 7261 TTTGGTTTGGATCAATAGGAAGCATATGAGCAACCAACCATGCAAAATGTTTGAAT 7320
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAGTACCCAAACACATTTCTGCTTTTCACTTAAAGTGT 7728
Db 7321 GATATGACCAAAATTTTAAAGTAGGAAGTACCCAAACACATTTCTGCTTTTCACTTAAAGTGT 7380
QY 7729 CTGGCCCGCATATCTGTAGGAACAGCATGATCTTTTACTGTGATATTTTAAATATCCA 7788
Db 7381 CTGGCCCGCATATCTGTAGGAACAGCATGATCTTTTACTGTGATATTTTAAATATCCA 7440
QY 7789 CAGTACT 7795
Db 7441 CAGTACT 7447

RESULT 10
AX277596 7680 bp DNA linear PAT 01-NOV-2001
LOCUS AX277596
DEFINITION Sequence 75 from Patent WO0177327.
ACCESSION AX277596
VERSION AX277596.1 GI:16604795
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Carulli,J.P., Little,R.D., Recker,R.R. and Johnson,M.L.
The high bone mass gene of Il13.3
Patent: WO 017327-A 75 18-OCT-2001;
Genome Therapeutics Corporation (US)
FEATURES
    source
        1..7680
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match 95.2%; Score 7423; DB 6; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 349 GAAGAGCAGAGGAGGCTCAGCAATGGTTGAGCCCGAGTCCCGGTGGCTGTCACTCA 408
Db 1 GAAGAGCAGAGGAGGCTCAGCAATGGTTGAGCCCGAGTCCCGGTGGCTGTCACTCA 60
QY 409 AAGCAAGCCCGGTTGTTATGACAAATGAAACACATATCAGATAAATCAACAGTGGAGCG 468
Db 61 AAGCAAGCCCGGTTGTTATGACAAATGAAACACATATCAGATAAATCAACAGTGGAGCG 120
QY 469 GACCTACTAGGCAATCGCTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 528
Db 121 GACCTACTAGGCAATCGCTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGCAGAGTACACTGGGAAACACTTACC 588
Db 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGCAGAGTACACTGGGAAACACTTACC 240
QY 589 AGTGGGTGACACTTATGAGCGGTCCTAAAGACTCCATGATCTGGGACTGTACTCTGATCGG 648
```

```
Db 241 AGTGGGTGACACTTATGAGCGCTCTAAAGACTCTCCATGATCTGGGACTGTACCTGCATCGG 300
QY 649 GGTCTGGCGAGGAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTTCAGTTC 708
Db 301 GGTCTGGCGAGGAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTTCAGTTC 360
QY 709 CTACAAGATTGGTGACACTCTGGAGAGACCAATGAGACTGGTGTGTATCATGTTAGAGTG 768
Db 361 CTACAAGATTGGTGACACTCTGGAGAGACCAATGAGACTGGTGTGTATCATGTTAGAGTG 420
QY 769 TGTGTGCTCTGTGTAATGGAAGAGGAGATGGAAGCTGCAAGCCCATAGCTGAGAAGTGT 828
Db 421 TGTGTGCTCTGTGTAATGGAAGAGGAGATGGAAGCTGCAAGCCCATAGCTGAGAAGTGT 480
QY 829 TGATCATGCTGCTGGGACTTCTTATGTGCTCGAGAAAAGCTGGGAGAGCCCTACCAAGG 888
Db 481 TGATCATGCTGCTGGGACTTCTTATGTGCTCGAGAAAAGCTGGGAGAGCCCTACCAAGG 540
QY 889 CTGGATGATGTTAGATTTGACTTCTGGGAGAGGAGCGGAGCGCATCATCTTGCACCTTC 948
Db 541 CTGGATGATGTTAGATTTGACTTCTGGGAGAGGAGCGGAGCGCATCATCTTGCACCTTC 600
QY 949 TAGAATATAGATCAACAGCATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAG 1008
Db 601 TAGAATATAGATCAACAGCATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAG 660
QY 1009 CAAGAAGGATATCGAGGAAACCTGCTCCAGTGTCTCTGACAGGCAAGCGGAGGAGGAGA 1068
Db 661 CAAGAAGGATATCGAGGAAACCTGCTCCAGTGTCTCTGACAGGCAAGCGGAGGAGGAGA 720
QY 1069 GTGGAAGTGTGAGAGGACACACCTCTGTGACAGACCATCTCGAGCGGATCTGGCCCTTCAC 1128
Db 721 GTGGAAGTGTGAGAGGACACACCTCTGTGACAGACCATCTCGAGCGGATCTGGCCCTTCAC 780
QY 1129 CGATGTTCTGTGAGCTGTTTCAACCCGAGCTCACCCCGAGCTCTCCCTATGGCCA 1188
Db 781 CGATGTTCTGTGAGCTGTTTCAACCCGAGCTCACCCCGAGCTCTCCCTATGGCCA 840
QY 1189 CTGTGTCAAGAGCAGTGTGTGCTTCTCTCTGGGGATGCAAGTGGCTGGAAGACACAAAG 1248
Db 841 CTGTGTCAAGAGCAGTGTGTGCTTCTCTCTGGGGATGCAAGTGGCTGGAAGACACAAAG 900
QY 1249 AATAAGCAAAATGCTTTGACGCTGTGGGCAACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1308
Db 901 AATAAGCAAAATGCTTTGACGCTGTGGGCAACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 1309 AACCCAGACTTACGGTGGCAAATCTCAAAATGGAGAGCATGTGTCTTACCATTCACCTACAA 1368
Db 961 AACCCAGACTTACGGTGGCAAATCTTAAATGGAGAGCATGTGTCTTACCATTCACCTACAA 1020
QY 1369 TGGCAGGAGCTTCTACTCTCTGACACAGAGGGGCAAGGAGGAGATCTTTGGTGGAG 1428
Db 1021 TGGCAGGAGCTTCTACTCTCTGACCAACGAGAGGGGCAAGGAGGAGATCTTTGGTGGAG 1080
QY 1429 CACAACTTCGAATATGAGCAGGAGCCAGAAATACTCTTTCTGCACAGACCACTGTTT 1488
Db 1081 CACAACTTCGAATATGAGCAGGAGCCAGAAATACTCTTTCTGCACAGACCACTGTTT 1140
QY 1489 GGTTCAGACTCGAGGAGGAAATTCCTAATGGTGTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCT 1548
Db 1141 GGTTCAGACTCAAGGAGGAAATTCCTAATGGTGTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCT 1200
QY 1549 CAACCACTTACACTGATTGCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1608
Db 1201 CAACCACTTACACTGATTGCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1609 GACCAACAGAACTATGATGCCGACAGAAAGTTTGGGTTTCTGCCCATGGCTGCCCAACA 1668
Db 1261 GACCAACAGAACTATGATGCCGACAGAAAGTTTGGGTTTCTGCCCATGGCTGCCCAACA 1320
QY 1669 GGAATCTGCACAACTATGAGGGGTCTATGTCGCTTGGAGTACCTGGGATGAGGATGAGGATGAGG 1728
```

1321 GGAATCTGCACACCAATGAAGGGGTCTGTACCGCATTTGGAGATCAGTGGGATAGCA 1380  
1729 GCATGACATGGGTGACATGATGAGGTGCGAGTGTGTGGAAATGGTCGTGGGAAATGGAC 1788  
1381 GCATGACATGGGTGACATGATGAGGTGCGAGTGTGTGGAAATGGTCGTGGGAAATGGAC 1440  
1789 ATGCAATTCCTACTCGCAGCTTCAGAGATCAGTGCATTTGTGATGACATCAGTATCAATGT 1848  
1441 ATGCAATTCCTACTCGCAGCTTCAGAGATCAGTGCATTTGTGATGACATCAGTATCAATGT 1500  
1849 GAACGACACATTCACAAAGCGTCAATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG 1908  
1501 GAACGACACATTCACAAAGCGTCAATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG 1560  
1909 TCAGGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968  
1561 TCAGGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620  
1969 GACGTTTATCAAAATGGAGATTCATGGGAGAAATGATGCAATGGTGTGATGATGCTTCCAG 2028  
1621 GACGTTTATCAAAATGGAGATTCATGGGAGAAATGATGCAATGGTGTGATGATGCTTCCAG 1680  
2029 CTACTGCTATGGCGGTGGCAATTCGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAG 2088  
1681 CTACTGCTATGGCGGTGGCAATTCGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAG 1740  
2089 CTCAGGTGTCCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACACC 2148  
1741 CTCAGGTGTCCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACACC 1800  
2149 CATCCAGTGAATGCACCAAGCATCTACATTTCCAAAGTACATTTCTCAGTGGAGACC 2208  
1801 CATCCAGTGAATGCACCAAGCATCTACATTTCCAAAGTACATTTCTCAGTGGAGACC 1860  
2209 TAAAAATTCGTAGGCGGTGGAGGAAGCTACATACAGAGGCACTTAACTCTCTACAC 2268  
1861 TAAAAATTCGTAGGCGGTGGAGGAAGCTACATACAGAGGCACTTAACTCTCTACAC 1920  
2269 CATCAAGGCTGAAAGCTGGTGTGATACAGAGGCGCAGCTCATCAGCATCCAGCAGTA 2328  
1921 CATCAAGGCTGAAAGCTGGTGTGATACAGAGGCGCAGCTCATCAGCATCCAGCAGTA 1980  
2329 CGGCACCAAGAGTGAATGCTGCTTGAATTCACACACACAGCAGCAGCAGCTGTGAC 2388  
1981 CGGCACCAAGAGTGAATGCTGCTTGAATTCACACACACAGCAGCAGCAGCTGTGAC 2040  
2389 CAGCAACACCGTGCAGAGGAGAGCAGTCCCTTTTCTCTTTGTCGTCAGTCTGAAATC 2448  
2041 CAGCAACACCGTGCAGAGGAGAGCAGTCCCTTTTCTCTTTGTCGTCAGTCTGAAATC 2100  
2449 TGTGACCGAATCACAGCCAGTACCTTTGTCGTCAGTCTGTCAGTCTGTCAGTCTGTCAG 2508  
2101 TGTGACCGAATCACAGCCAGTACCTTTGTCGTCAGTCTGTCAGTCTGTCAGTCTGTCAG 2160  
2509 GTCCGGAATCCGGGTGGAATATGAGCTGAGTGAAGGAGGAGATGAGCCACAGTACCTGGA 2568  
2161 GTCCGGAATCCGGGTGGAATATGAGCTGAGTGAAGGAGGAGATGAGCCACAGTACCTGGA 2220  
2569 TCTTCCAGCAGCAGCAGTCTGTCGTAACCTCCTGAGCTGCTTCTGTCGTCGTCGTCGTCG 2628  
2221 TCTTCCAGCAGCAGCAGTCTGTCGTAACCTCCTGAGCTGCTTCTGTCGTCGTCGTCGTCG 2280  
2629 TGTAAATGCTATCAGATATCTGAGGATGGGAGCAGAGTTTGCATCCTGCTCTACTTCA 2688  
2281 TGTAAATGCTATCAGATATCTGAGGATGGGAGCAGAGTTTGCATCCTGCTCTACTTCA 2340  
2689 AACCAAGCGCTGATGCGCTCTGACCCGCTGACGCTGTCGTCGTCGTCGTCGTCGTCGTCG 2748  
2341 AACCAAGCGCTGATGCGCTCTGACCCGCTGACGCTGTCGTCGTCGTCGTCGTCGTCGTCG 2400  
2749 TGTGTTTCGCTGGAGCAGACCCAGGCTCCCATCAGAGGTACAGAAATAGTCTATTTCGCC 2808  
2401 TGTGTTTCGCTGGAGCAGACCCAGGCTCCCATCAGAGGTACAGAAATAGTCTATTTCGCC 2460

2809 ATCAGTAGAAGGTAGCAGCAAGAACTCAACCTTCTCTGAAACTCTGCAAACTCCGTCACCCCT 2868  
2461 ATCAGTAGAAGGTAGCAGCAAGAACTCAACCTTCTCTGAAACTCTGCAAACTCCGTCACCCCT 2520  
2869 CAGTGAATTCGCAACTGCTGTTAGTATAACATCACTATCTATCTATCTGCTGGAAGAAATCA 2928  
2521 CAGTGAATTCGCAACTGCTGTTAGTATAACATCACTATCTATCTATCTGCTGGAAGAAATCA 2580  
2929 AGAAAGTACACCTGTTGTCATTAACAAGAAACCACTGGGACCCCAAGCTCAGATACAGT 2988  
2581 AGAAAGTACACCTGTTGTCATTAACAAGAAACCACTGGGACCCCAAGCTCAGATACAGT 2640  
2989 GCGCTCTCCAGGAGACTGCGAGTTTGTGGAAGTACAGAGCTGAAGGTCAACATCATGTG 3048  
2641 GCGCTCTCCAGGAGACTGCGAGTTTGTGGAAGTACAGAGCTGAAGGTCAACATCATGTG 2700  
3049 GACACCGCTCAGAGTGCAGTACCGGTACCGGTGATGATGATGATGATGATGATGATGATGATG 3108  
2701 GACACCGCTCAGAGTGCAGTACCGGTACCGGTGATGATGATGATGATGATGATGATGATGATG 2760  
3109 TGGGAGCAGCGGCGAGAGGCTGCCCATCAGCAGGAACACCTTTGTCAGAGGTCAACCGGCT 3168  
2761 TGGGAGCAGCGGCGAGAGGCTGCCCATCAGCAGGAACACCTTTGTCAGAGGTCAACCGGCT 2820  
3169 GTCCCTCGGGGTCACTATTACTTTCAAAGTCTTTTGCAGTGCAGCCATGGGAGGAGAGCAA 3228  
2821 GTCCCTCGGGGTCACTATTACTTTCAAAGTCTTTTGCAGTGCAGCCATGGGAGGAGAGCAA 2880  
3229 GCGCTGACTGCTCAACAGACAACTGGATGCTCCCACTCACTCACTCACTCACTCACTCACTCA 3288  
2881 GCGCTGACTGCTCAACAGACAACTGGATGCTCCCACTCACTCACTCACTCACTCACTCACTCA 2940  
3289 TGAAGTCAATTTCTACTGCTGTTGAGATGAGCTCCACCTCGGGCCAGATPACAGGATA 3348  
2941 TGAAGTCAATTTCTACTGCTGTTGAGATGAGCTCCACCTCGGGCCAGATPACAGGATA 3000  
3349 CCGACTGACCTGGGCTTACCCGAGAGGAGACCCAGGAGTACATGATGATGATGATGATGATGATG 3408  
3001 CCGACTGACCTGGGCTTACCCGAGAGGAGACCCAGGAGTACATGATGATGATGATGATGATGATG 3060  
3409 TGTCTCAAGTACCCACTGAGGAATCTGAGGCTGCACTGAGTACACCGCTATCCCTCGT 3468  
3061 TGTCTCAAGTACCCACTGAGGAATCTGAGGCTGCACTGAGTACACCGCTATCCCTCGT 3120  
3469 GGCATTAAGGGGCAACAGAGAGCCCAAGCCACTGAGTCTTTTACCACTGCGAGCC 3528  
3121 GGCATTAAGGGGCAACAGAGAGCCCAAGCCACTGAGTCTTTTACCACTGCGAGCC 3180  
3529 TGGGAGCTCTATTCCACTTACACCGAGTCACTGAGACCACTGATGATGATGATGATGATGATG 3588  
3181 TGGGAGCTCTATTCCACTTACACCGAGTCACTGAGACCACTGATGATGATGATGATGATGATG 3240  
3589 GACGCTGCTCCAAAGAAATGGTTTAAAGTGGGTGTAACGACCAAGCCAGGAGGAGAGCC 3648  
3241 GACGCTGCTCCAAAGAAATGGTTTAAAGTGGGTGTAACGACCAAGCCAGGAGGAGAGCC 3300  
3649 ACCAGGAGTGAATTCAGACTCAGGAGCATGTTGTCGCGCTTGAATCCAGAGT 3708  
3301 ACCAGGAGTGAATTCAGACTCAGGAGCATGTTGTCGCGCTTGAATCCAGAGT 3360  
3709 AGAATAGCTCTACACCACTCCAAAGTCTGAGAGATGGAACAGAAAGAGATGCGCAATTTGT 3768  
3361 AGAATAGCTCTACACCACTCCAAAGTCTGAGAGATGGAACAGAAAGAGATGCGCAATTTGT 3420  
3769 AAAAAGAGTGTGACCACTTCTCCCAACCAACTTGCATCTGAGGAGCAACCCCTGA 3828  
3421 AAAAAGAGTGTGACCACTTCTCCCAACCAACTTGCATCTGAGGAGCAACCCCTGA 3480  
3829 CACTGGAGTGTCTCAGTCTCTCGGAGAGAGGAGCACCACCCAGACATTACTGTTTATAG 3888  
3481 CACTGGAGTGTCTCAGTCTCTCGGAGAGAGGAGCACCACCCAGACATTACTGTTTATAG 3540

3889 AATTACCAACCCCTTAACAACGGCCAGAGGAAATCTTTTGGAGAGTGGTCCATGC 3948  
3541 AATTACCAACCCCTTAACAACGGCCAGAGGAAATCTTTTGGAGAGTGGTCCATGC 3600  
3949 TGATCAGAGCTCCTGCACCTTTTGAATAAAGTCCCGGCTGGAGTACAAATGTCAAGTGT 4008  
3601 TGATCAGAGCTCCTGCACCTTTTGAATAAAGTCCCGGCTGGAGTACAAATGTCAAGTGT 3660  
4009 TTACACTGTCAAGGATGACAAAGGAGTGCCTTAATCTCTGATACCAATCATCCAGCTGT 4068  
3661 TTACACTGTCAAGGATGACAAAGGAGTGCCTTAATCTCTGATACCAATCATCCAGCTGT 3720  
4069 TCCTCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACACCATGGCTGTCACTG 4128  
3721 TCCTCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACACCATGGCTGTCACTG 3780  
4129 GGCTCCACCCCATCATTTGATTTAACCAATTTCTGCTGGTGTACTCACTGTGAAAAA 4188  
3781 GGCTCCACCCCATCATTTGATTTAACCAATTTCTGCTGGTGTACTCACTGTGAAAAA 3840  
4189 TGAGGAGATGTTGAGAGTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAACAAA 4248  
3841 TGAGGAGATGTTGAGAGTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAACAAA 3900  
4249 TCTCTGCTGTTACAGATATGTAGTGTCTCTCCAGTGTCTACGAAACCAATGAGAG 4308  
3901 TCTCTGCTGTTACAGATATGTAGTGTCTCTCCAGTGTCTACGAAACCAATGAGAG 3960  
4309 CACACTCTTAGAGGAGACAGAAACAGTCTTGAATCCCACTGCGATGTCTTTTC 4368  
3961 CACACTCTTAGAGGAGACAGAAACAGTCTTGAATCCCACTGCGATGTCTTTTC 4020  
4369 TGATATTACTGCCAATCTTTTACTGTGACCTGAGTGTCTCTCGAGCCACCATCTGG 4428  
4021 TGATATTACTGCCAATCTTTTACTGTGACCTGAGTGTCTCTCGAGCCACCATCTGG 4080  
4429 CTACAGGATCCGCAATCATCCGAGCACTTACGTGGGAGACCTCGAGAAAGATCGGTGCC 4488  
4081 CTACAGGATCCGCAATCATCCGAGCACTTACGTGGGAGACCTCGAGAAAGATCGGTGCC 4140  
4489 CCACTCTCGGAATTCATCACTCCTCAACCTCTCACTCCAGGACAGAGTATGTGGTCAG 4548  
4141 CCACTCTCGGAATTCATCACTCCTCAACCTCTCACTCCAGGACAGAGTATGTGGTCAG 4200  
4549 CATGTGTCTTTAATGGCAGAGAGAAAGTCCCTTTATTTGATGGCCAAATCAACAGT 4608  
4201 CATGTGTCTTTAATGGCAGAGAGAAAGTCCCTTTATTTGATGGCCAAATCAACAGT 4260  
4609 TTCTGATGTTCCGAGGACCTGGAAGTGTCTGCGACCCCTCACTGCTACTGATCATCAG 4668  
4261 TTCTGATGTTCCGAGGACCTGGAAGTGTCTGCGACCCCTCACTGCTACTGATCATCAG 4320  
4669 CTGGGATGCTCCTGTCTGTCACAGTGAGATATTACAGGATCACTTACGAGAAACAGGAGG 4728  
4321 CTGGGATGCTCCTGTCTGTCACAGTGAGATATTACAGGATCACTTACGAGAAACAGGAGG 4380  
4729 AANTAGCCTGTCCAGAGTCTCACTGCTGGGAGCAAGTCTACAGTACCATCAGCGG 4788  
4381 AANTAGCCTGTCCAGAGTCTCACTGCTGGGAGCAAGTCTACAGTACCATCAGCGG 4440  
4789 CCTTAAACCTGGAGTGTGATTTACCACTCACTGTGTATGCTGCTCACTGGCCGTGGAGACAG 4848  
4441 CCTTAAACCTGGAGTGTGATTTACCACTCACTGTGTATGCTGCTCACTGGCCGTGGAGACAG 4500  
4849 CCCGGCAGCAGCAAGCAATTTCCATTAATTTACCGAAGAGAAATTTGACCAACCATCCCA 4908  
4501 CCCGGCAGCAGCAAGCAATTTCCATTAATTTACCGAAGAGAAATTTGACCAACCATCCCA 4560  
4909 GATGCAAGTCAAGGATGTTTACGAGCAACAGCAATTTAGTGTCAAGTGGCTCCCTTCAAGTTC 4968  
4561 GATGCAAGTCAAGGATGTTTACGAGCAACAGCAATTTAGTGTCAAGTGGCTCCCTTCAAGTTC 4620  
4969 CCTGTGTTACAGAGTAAACCACTCCCAAAAATTTGACCAACCATCCCAACCAAAAC 5028

4621 CCTGTGTTACTGGTTACAGAGTAAACCACTCCCAAAAATTTGACCAACCAACCAAAAC 4680  
5029 TAAACTGCAAGTCCAGATCAAAACAGAAATGACTATTGAAGCTTTGACGCCACAGTGA 5088  
4681 TAAACTGCAAGTCCAGATCAAAACAGAAATGACTATTGAAGCTTTGACGCCACAGTGA 4740  
5089 GTATGTGGTGTAGTGTCTGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTTCTGGTTCAGAC 5148  
4741 GTATGTGGTGTAGTGTCTGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTTCTGGTTCAGAC 4800  
5149 TGCAAGTAAACCAATGATCGCCCTAAAGGACTGCGCAATTCATGTATGTGGATGTGATTC 5208  
4801 TGCAAGTAAACCAATGATCGCCCTAAAGGACTGCGCAATTCATGTATGTGGATGTGATTC 4860  
5209 CATCAAAATTTGCTTGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTTCTC 5268  
4861 CATCAAAATTTGCTTGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTTCTC 4920  
5269 GAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGCACTGATGTGTAAGAGACACTGC 5328  
4921 GAGCCCTGAGGATGGAATCCATGAGCTATTTCCTGCACTGATGTGTAAGAGACACTGC 4980  
5329 AGAGCTGCAAGGCTCAGACCCGGTTCAGTACACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5388  
4981 AGAGCTGCAAGGCTCAGACCCGGTTCAGTACACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5040  
5389 TGATATTGAGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTTCCTGCAACACTGA 5448  
5041 TGATATTGAGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTTCCTGCAACACTGA 5100  
5449 CTTGAAAGTTCATCAGGTCAACCCCAAGCCCTGAGCCGCTGAGTCAAGTCAAGTCAAGTCAAGT 5508  
5101 CTTGAAAGTTCATCAGGTCAACCCCAAGCCCTGAGCCGCTGAGTCAAGTCAAGTCAAGTCAAGT 5160  
5509 TCAGCTCAGTGGATATCGAGTGGGTGACCCCAAGGAGAGACCCGACCAATGAAGA 5568  
5161 TCAGCTCAGTGGATATCGAGTGGGTGACCCCAAGGAGAGACCCGACCAATGAAGA 5220  
5569 AATCAACCTTGTCTCTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAA 5628  
5221 AATCAACCTTGTCTCTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAA 5280  
5629 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTTGAAGGAGAGTCAAGTCAAGTCAAGT 5688  
5281 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTTGAAGGAGAGTCAAGTCAAGTCAAGT 5340  
5689 TGTCAACCTCTCTGGAGATGTCAAGCCCAAGGAGTCTGTTGACAGATGCTACTGA 5748  
5341 TGTCAACCTCTCTGGAGATGTCAAGCCCAAGGAGTCTGTTGACAGATGCTACTGA 5400  
5749 GACCAATCAGTGTAGTGGAGAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5808  
5401 GACCAATCAGTGTAGTGGAGAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5460  
5809 TGCCGTTTCCAGCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAAGAG 5868  
5461 TGCCGTTTCCAGCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAAGAG 5520  
5869 CTACACCATCAGTGGCTTACCAACAGGCACTGACTCAAGATCTACTGTACACCTTGA 5928  
5521 CTACACCATCAGTGGCTTACCAACAGGCACTGACTCAAGATCTACTGTACACCTTGA 5580  
5929 TGCAATGCTCGGAGTCTCCCTGTGTCTGATCGAGCCCTCCACTGCCATTTGATGCCAATC 5988  
5581 TGCAATGCTCGGAGTCTCCCTGTGTCTGATCGAGCCCTCCACTGCCATTTGATGCCAATC 5640  
5989 CAACCTCGGTTTCTGGCCAGCACCCCAATTTCTGCTGGTATCATGCGAGCCGACG 6048  
5641 CAACCTCGGTTTCTGGCCAGCACCCCAATTTCTGCTGGTATCATGCGAGCCGACG 5700  
6049 TGCCAGGATTTACCGGCTTACATCATCAAGTATGAGAAAGCTGCTCTCTCCCAAGAGT 6108





Matches	7432;	Conservative	0;	Mismatches	15;	Indels	0;	Gaps	0;
QY	349	GAAGAGCAAGAGCAGGCTCAGCAAATGGTTACGCCCAAGTCCCCGGTGGCTGTCA	GTCA	408					
Db	1	GAAGAGCAAGAGCAGGCTCAGCAAATGGTTACGCCCAAGTCCCCGGTGGCTGTCA	GTCA	60					
QY	409	AAGCAAGCCCGGTGTTATGACAATGGAAAACTATCAGATAAACTCAACAGTGGGAGG	G	468					
Db	61	AAGCAAGCCCGGTGTTATGACAATGGAAAACTATCAGATAAACTCAACAGTGGGAGG	G	120					
QY	469	GACCTACTAGGCAATGCGTGTGTTGTAAGTCTTATGGAGGAAGCCGAGTTTAACTG	528						
Db	121	GACCTACTAGTAAATGTGTGTGTTGTAAGTCTTATGGAGGAAGCCGAGTTTAACTG	180						
QY	529	CGAAAGTAAACCTGAAGCTCAAGAGACTTGCTTTGACAACTACACTGGGAACTTACCG	588						
Db	181	CGAAAGTAAACCTGAAGCTCAAGAGACTTGCTTTGACAACTACACTGGGAACTTACCG	240						
QY	589	AGTGGGTGACATTTATGAGCGTCTTAAAGTCCATGATCTGGGATGTACCTGCATCGG	648						
Db	241	AGTGGGTGACATTTATGAGCGTCTTAAAGTCCATGATCTGGGATGTACCTGCATCGG	300						
QY	649	GGCTGGCGAGGAGAAATAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAGTC	708						
Db	301	GGCTGGCGAGGAGAAATAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAGTC	360						
QY	709	CTACAAGATTGGTGAACCTGGAGGAGACCAATGAGACTGTGGTGGTTACATGTTAGAGTG	768						
Db	361	CTACAAGATTGGTGAACCTGGAGGAGACCAATGAGACTGTGGTGGTTACATGTTAGAGTG	420						
QY	769	TGTGTGTTCTGTTAAATGAAAAAGAGAAATCGAAGCTGCAAGCCCATAGCTGAGAAGTGT	828						
Db	421	TGTGTGTTCTGTTAAATGAAAAAGAGAAATCGAAGCTGCAAGCCCATAGCTGAGAAGTGT	480						
QY	829	TGATCATGCTGCTGGGACTTCTATGTGTGTGGAGAAAAGCTGGGAGAGCCCTACCAAG	888						
Db	481	TGATCATGCTGCTGGGACTTCTATGTGTGTGGAGAAAAGCTGGGAGAGCCCTACCAAG	540						
QY	889	CTGGATGATGGTATGATGTAATCTTCCTCTGGGAGAAAGCAGCGGACCGCATCACTTGCAC	948						
Db	541	CTGGATGATGGTATGATGTAATCTTCCTCTGGGAGAAAGCAGCGGACCGCATCACTTGCAC	600						
QY	949	TAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAAATTTGGAGACACCTGGAG	1008						
Db	601	TAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAAATTTGGAGACACCTGGAG	660						
QY	1009	CAAGAAGGATAATCGAGGAAAACCTGCTTCAGTGTGATCTGCACAGGCAACGGCCGAGGAGA	1068						
Db	661	CAAGAAGGATAATCGAGGAAAACCTGCTTCAGTGTGATCTGCACAGGCAACGGCCGAGGAGA	720						
QY	1069	GTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAATCGAGGGGATCTGGCCCTTCA	1128						
Db	721	GTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAATCGAGGGGATCTGGCCCTTCA	780						
QY	1129	CGATGTTTCGTGACGCTGTTTACCAACCGCAGCCTCACCCCCAGCCTCTCCCTATGGCCA	1188						
Db	781	CGATGTTTCGTGACGCTGTTTACCAACCGCAGCCTCACCCCCAGCCTCTCCCTATGGCCA	840						
QY	1189	CTGTGTCAACAGACAGTGGTGTGTTACTCTGTGGGATGCGAGTGGCTGAAGACACAAGG	1248						
Db	841	CTGTGTCAACAGACAGTGGTGTGTTACTCTGTGGGATGCGAGTGGCTGAAGACACAAGG	900						
QY	1249	AAATAAGCAAAATGCTTTTGACGCTGCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT	1308						
Db	901	AAATAAGCAAAATGCTTTTGACGCTGCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT	960						
QY	1309	AACCCGACCTTACGGTGGCACTCAAAATGGAGAGCCATGTCTTACCATTCACCTACAA	1368						
Db	961	AACCCGACCTTACGGTGGCACTTAAATGGAGAGCCATGTCTTACCATTCACCTACAA	1020						
QY	1369	TGGCAGGACGTTCTACTCTGCAACCAAGAGGGCGACAGGACGACATCTTTGGTGCAG	1428						
Db	1021	TGGCAGGACGTTCTACTCTGCAACCAAGAGGGCGACAGGACGACATCTTTGGTGCAG	1080						

Qy	1429	CACAACTTCGAATTATGACGAGGACGAGAAATACTCTTCTGTGCACAGACCACTACTGTTTT	1488
Db	1081	CACAACTTCGAATTATGACGAGGACGAGAAATACTCTTCTGTGCACAGACCACTACTGTTTT	1140
Qy	1489	GGTTTCAGACTCGAGGAGGAAATTCCAATGGTGGCTTGTGCACCTTCCCCTTCTCTATACAA	1548
Db	1141	GGTTTCAGACTCGAGGAGGAAATTCCAATGGTGGCTTGTGCACCTTCCCCTTCTCTATACAA	1200
Qy	1549	CAACCACAATTAACATGATTGCACTTCTGAGGGCAGAGAGACCAATGAAAGTGGTGTGG	1608
Db	1201	CAACCACAATTAACATGATTGCACTTCTGAGGGCAGAGAGACCAATGAAAGTGGTGTGG	1260
Qy	1609	GACCAACAGAACTATGATGCGCGACACAGAAATTTGGTTCTGCCCCATGGCTGCCACGA	1668
Db	1261	GACCAACAGAACTATGATGCGCGACACAGAAATTTGGTTCTGCCCCATGGCTGCCACGA	1320
Qy	1669	GGAAATTCGCAACCAATGAAGGGTCACTATACCGCATTTGGAGATCACTGGGATGAACA	1728
Db	1321	GGAAATTCGCAACCAATGAAGGGTCACTATACCGCATTTGGAGATCACTGGGATGAACA	1380
Qy	1729	GCATGACATGGGTACATGATGAGGTGCACGTGTGTTGGGAAATGGCTGGGGAAATGGAC	1788
Db	1381	GCATGACATGGGTACATGATGAGGTGCACGTGTGTTGGGAAATGGCTGGGGAAATGGAC	1440
Qy	1789	ATGCATATGCCCTACTCGAGCTTCGAGATCAGTGCATTTGTGATGACATCACTTACAAATGT	1848
Db	1441	ATGCATATGCCCTACTCGAGCTTCGAGATCAGTGCATTTGTGATGACATCACTTACAAATGT	1500
Qy	1849	GAAACGACACATTCACAAGCGTCAATGAAGAGGGGCACATGCTGAACTGTACATGCTTCGG	1908
Db	1501	GAAACGACACATTCACAAGCGTCAATGAAGAGGGGCACATGCTGAACTGTACATGCTTCGG	1560
Qy	1909	TCAGGTGGGGGACAGGTGGAAGTGTGATCCCTGCGCAATGCGAGGATTCAGAGACTGG	1968
Db	1561	TCAGGTGGGGGACAGGTGGAAGTGTGATCCCTGCGCAATGCGAGGATTCAGAGACTGG	1620
Qy	1969	GACGTTTTATCAAAATGGAGATTCATGGGAGAAAGTATGTCATCGTGTGCAGATCCAGTG	2028
Db	1621	GACGTTTTATCAAAATGGAGATTCATGGGAGAAAGTATGTCATCGTGTGCAGATCCAGTG	1680
Qy	2029	CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATGCCAACCTTACAGACCTATCCAAG	2088
Db	1681	CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATGCCAACCTTACAGACCTATCCAAG	1740
Qy	2089	CTCAAGTGGTCTGTGGAATTTATCACTGAGACTCCGAGTCAGCCCAACTCCCAACC	2148
Db	1741	CTCAAGTGGTCTGTGGAATTTATCACTGAGACTCCGAGTCAGCCCAACTCCCAACC	1800
Qy	2149	CATCCAGTGAATGCACACAGCATCTACATTTCCAAAGTACATTTCTCAGGTGGAGACC	2208
Db	1801	CATCCAGTGAATGCACACAGCATCTACATTTCCAAAGTACATTTCTCAGGTGGAGACC	1860
Qy	2209	TAAAAATTCGTAGCCGTTTGGAGGAAGCTACCATACAGGGCCACTTAAACTCCTACAC	2268
Db	1861	TAAAAATTCGTAGCCGTTTGGAGGAAGCTACCATACAGGGCCACTTAAACTCCTACAC	1920
Qy	2269	CATCAAAAGGCTGNAAGCTGGTGTGGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA	2328
Db	1921	CATCAAAAGGCTGNAAGCTGGTGTGGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTA	1980
Qy	2329	CGGCCACCAAGAAGTGACTCGCTTTTGACTTTCACACCAACGAGCACCAAGCACTGTGAC	2388
Db	1981	CGGCCACCAAGAAGTGACTCGCTTTTGACTTTCACACCAACGAGCACCAAGCACTGTGAC	2040
Qy	2389	CAGCAACACCGTGCACGAGAGAGCACTCCCTTTTCTCCTTGTGGCCACTTCTGAAATC	2448
Db	2041	CAGCAACACCGTGCACGAGAGAGCACTCCCTTTTCTCCTTGTGGCCACTTCTGAAATC	2100
Qy	2449	TGTGACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTGGGTCTACGCTCCGACACCGT	2508
Db	2101	TGTGACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTGGGTCTACGCTCCGACACCGT	2160





Db 4321 CTGGAGTCTCTCTGTACAGTGAATATTACAGGATCACTTACGGAGAAACAGGAGG 4380  
Qy 4729 AAATAGCCCTGTCCAGAGTTTCACTGTGCTGGAGCAAGTCTACAGCTACCAATCAGCGG 4788  
Db 4381 AAATAGCCCTGTCCAGAGTTTCACTGTGCTGGAGCAAGTCTACAGCTACCAATCAGCGG 4440  
Qy 4789 CCTTAAACCTGGAGTTGATATACCATCATCTGTGTATGCTGTCTACTGGCCGTGGAGACAG 4848  
Db 4441 CCTTAAACCTGGAGTTGATATACCATCATCTGTGTATGCTGTCTACTGGCCGTGGAGACAG 4500  
Qy 4849 CCCGCAAGCAGCAAGCAATTTCCATTAATTAACGACAGAAATTCACAAAACATCCCA 4908  
Db 4501 CCCGCAAGCAGCAAGCAATTTCCATTAATTAACGACAGAAATTCACAAAACATCCCA 4560  
Qy 4909 GATGCAAGTGAACGATGTTTACGACCAACAGCATTAAGTGTCAAGTGGCTGCTTCAAGTTT 4968  
Db 4561 GATGCAAGTGAACGATGTTTACGACCAACAGCATTAAGTGTCAAGTGGCTGCTTCAAGTTT 4620  
Qy 4969 CCCTGTTTACTGGTTACAGAGTAAACCAACCTCCCAAAATGACAGGACCAACAAAAC 5028  
Db 4621 CCCTGTTTACTGGTTACAGAGTAAACCAACCTCCCAAAATGACAGGACCAACAAAAC 4680  
Qy 5029 TAAACTGCAAGTCAAGATCAAAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 5088  
Db 4681 TAAACTGCAAGTCAAGATCAAAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 4740  
Qy 5089 GTATGTGTTAGTGTCTATGCTCAGATCCAAAGCGAGAGTCAAGCTCTGGTTACAG 5148  
Db 4741 GTATGTGTTAGTGTCTATGCTCAGATCCAAAGCGAGAGTCAAGCTCTGGTTACAG 4800  
Qy 5149 TGCAGTAAACCAACATGATGCCCTTAAAGGACTGGCATTTCACTGATGTGATGTCGATTC 5208  
Db 4801 TGCAGTAAACCAACATGATGCCCTTAAAGGACTGGCATTTCACTGATGTGATGTCGATTC 4860  
Qy 5209 CATCAAAATGCTTGGAAAGCCCAACAGGCGAAGTTCCAGGTACAGGTTGACTACTC 5268  
Db 4861 CATCAAAATGCTTGGAAAGCCCAACAGGCGAAGTTTCCAGGTACAGGTTGACTACTC 4920  
Qy 5269 GAGCCCTCAGGATGGAATCCATGAGCTATTCCTGCACTGATGCTGGAAGAACACTGC 5328  
Db 4921 GAGCCCTCAGGATGGAATCCATGAGCTATTCCTGCACTGATGCTGGAAGAACACTGC 4980  
Qy 5329 AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTGCTTGCCTTGACGA 5388  
Db 4981 AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTGCTTGCCTTGACGA 5040  
Qy 5389 TGATATGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTGCAACCACTGA 5448  
Db 5041 TGATATGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTGCAACCACTGA 5100  
Qy 5449 CTTGAAGTTTCACTAGGTCACACCAAGCCTGAGCCCTGAGCCAGTCCACACCAATGT 5508  
Db 5101 CTTGAAGTTTCACTAGGTCACACCAAGCCTGAGCCCTGAGCCAGTCCACACCAATGT 5160  
Qy 5509 TCAGTCTCACTGGATATCAGTCCGGTTCACCCCAAGAGAGACCCGACCAATGAAGA 5568  
Db 5161 TCAGTCTCACTGGATATCAGTCCGGTTCACCCCAAGAGAGACCCGACCAATGAAGA 5220  
Qy 5569 AATCAACCTTGTCTGTGACAGCTCATTCGGTGTGTTATACAGACTTATGGTGGCCACCA 5628  
Db 5221 AATCAACCTTGTCTGTGACAGCTCATTCGGTGTGTTATACAGACTTATGGTGGCCACCA 5280  
Qy 5629 ATATGAAGTGTGTTCTATGCTTTAAGGACACTTTGACAGCAGACCCAGCTCAGGGAGT 5688  
Db 5281 ATATGAAGTGTGTTCTATGCTTTAAGGACACTTTGACAGCAGACCCAGCTCAGGGTGT 5340  
Qy 5689 TGTCAACCTCTGGAGAAATGTCAGGCCCAAGAAAGGGCTGTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCAACCTCTGGAGAAATGTCAGGCCCAAGAAAGGGCTGTGTGACAGATGCTACTGA 5400  
Qy 5749 GACCACCATCAGATGAGTGGAGAACCAAGTCACTGGCCCTCCAGTTGA 5808  
Db 5401 GACCACCATCAGATGAGTGGAGAACCAAGTCACTGGCCCTCCAGTTGA 5460

Qy 5809 TGCCCTTCCAGCCCAATGGCCAGATCTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5868  
Db 5461 TGCCCTTCCAGCCCAATGGCCAGATCTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5520  
Qy 5869 CTACACCACTACTGGCTTACAAACAGGCACTGACTTACAGATCTTACCTGTGTACACCTTGA 5928  
Db 5521 CTACACCACTACTGGCTTACAAACAGGCACTGACTTACAGATCTTACCTGTGTACACCTTGA 5580  
Qy 5929 TGCAATGCTGGAGCTCCCTGTGGTCAATCAAGCTTCCAGCTTCCAGCTTGTGACCAATC 5988  
Db 5581 TGCAATGCTGGAGCTCCCTGTGGTCAATCAAGCTTCCAGCTTGTGACCAATC 5640  
Qy 5989 CAACCTGTGGTTTCTGGCCACCAACCCCAATTTCTTGTGTGTATCATGGCAGCCGCCACG 6048  
Db 5641 CAACCTGTGGTTTCTGGCCACCAACCCCAATTTCTTGTGTGTATCATGGCAGCCGCCACG 5700  
Qy 6049 TGCCAGGATTAACGGCTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGT 6108  
Db 5701 TGCCAGGATTAACGGCTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGT 5760  
Qy 6109 GGTCCCTCGGCCCGCTGTGTTCACAGAGCTACTATTCTGGCTTGGAAACCGGGAAC 6168  
Db 5761 GGTCCCTCGGCCCGCTGTGTTCACAGAGCTACTATTCTGGCTTGGAAACCGGGAAC 5820  
Qy 6169 CGAATATCAATTTATGTCAATTTGCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTTGG 6228  
Db 5821 CGAATATCAATTTATGTCAATTTGCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTTGG 5880  
Qy 6229 AAGGAAAGAGCAGAGAGCTTCCCACTGTAACTTCCACACCCCAATCTTCTCATGG 6288  
Db 5881 AAGGAAAGAGCAGAGAGCTTCCCACTGTAACTTCCACACCCCAATCTTCTCATGG 5940  
Qy 6289 ACCAGAGATCTTGGATGTTTCCCTTCCAGAGTCCAAAGACCCCTTTGTCACCAACCTGG 6348  
Db 5941 ACCAGAGATCTTGGATGTTTCCCTTCCAGAGTCCAAAGACCCCTTTGTCACCAACCTGG 6000  
Qy 6349 GTATGACATCTGAATGTTTACGCTTCTTCCAGCTTCCAAAGACCCCTTTGTCACCAACCTGG 6408  
Db 6001 GTATGACATCTGAATGTTTACGCTTCTTCCAGCTTCCAAAGACCCCTTTGTCACCAACCTGG 6060  
Qy 6409 GCAACCAATGATCTTTGAGGAACTGGTTTAAAGCGGACCAACCCGCCCAACAGCCCAAC 6468  
Db 6061 GCAACCAATGATCTTTGAGGAACTGGTTTAAAGCGGACCAACCCGCCCAACAGCCCAAC 6120  
Qy 6469 CCCCATAGGCTATAGCCCAAGACCATACCCCGGAGTGTAGGACCAAGAGCTCTCTCTCA 6528  
Db 6121 CCCCATAGGCTATAGCCCAAGACCATACCCCGGAGTGTAGGACCAAGAGCTCTCTCTCA 6180  
Qy 6529 GACCAACATCTCATGGGCCCTTCCAGGACACTTCTGAGTACATCATTTTCTATGTCATCC 6588  
Db 6181 GACCAACATCTCATGGGCCCTTCCAGGACACTTCTGAGTACATCATTTTCTATGTCATCC 6240  
Qy 6589 TGTGGCACTGATGAAGAACCTTACAGTTTCTGGGTTCTTGGAACTTCTTACAGTGGCCAC 6648  
Db 6241 TGTGGCACTGATGAAGAACCTTACAGTTTCTGGGTTCTTGGAACTTCTTACAGTGGCCAC 6300  
Qy 6649 TCTGACAGGCTCACCAGAGGTGCCACTTACCAATCATATAGTGGAGGCACTGAAAGACCA 6708  
Db 6301 TCTGACAGGCTCACCAGAGGTGCCACTTACCAATCATATAGTGGAGGCACTGAAAGACCA 6360  
Qy 6709 GCAGAGGCTAAGGTTTGGGAGAGGTTTACCGTGGGCAACTCTGTCAAAGAGGCTT 6768  
Db 6361 GCAGAGGCTAAGGTTTGGGAGAGGTTTACCGTGGGCAACTCTGTCAAAGAGGCTT 6420  
Qy 6769 GAACCAACCTACGGATGACTTCTGCTTTCAGCCCTACAGAGTGTCCCATTTAGCCCTTGG 6828  
Db 6421 GAACCAACCTACGGATGACTTCTGCTTTCAGCCCTACAGAGTGTCCCATTTAGCCCTTGG 6480  
Qy 6829 AGATGAGTGGGAAAGAAATGTCGAATCAGGCTTTAACTGTTGTGTCAGTGTGCTTGGCTT 6888  
Db 6481 AGATGAGTGGGAAAGAAATGTCGAATCAGGCTTTAACTGTTGTGTCAGTGTGCTTGGCTT 6540

[illegible]

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 7680)  
Kornblith,A.R., Vibe-Pedersen,K. and Baralle,F.E.  
Isolation and characterization of cDNA clones for human and bovine fibronectins  
Proc. Natl. Acad. Sci. U.S.A. 80 (11), 3218-3222 (1983)  
83221567  
6304699  
2 (bases 1 to 7680)  
Kornblith,A.R., Vibe-Pedersen,K. and Baralle,F.E.  
Human fibronectin: molecular cloning evidence for two mRNA species differing by an internal segment coding for a structural domain  
EMBO J. 3 (1), 221-226 (1984)  
84158533  
6200322  
3 (bases 1 to 7680)  
Kornblith,A.R., Vibe-Pedersen,K. and Baralle,F.E.  
Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats  
Nucleic Acids Res. 12 (14), 5853-5868 (1984)  
84272258  
6462919  
4 (bases 1 to 7680)  
Kornblith,A.R., Umazawa,K., Vibe-Pedersen,K. and Baralle,F.E.  
Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene  
EMBO J. 4 (7), 1755-1759 (1985)  
85284965  
2992939

Location/Qualifiers  
1..7680  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
<!.6988  
/codon\_start=2  
/product="fibronectin precursor"  
/protein\_id="CAA26536.1"  
/db\_xref="GI:31397"  
/db\_xref="SWISS-PROT:P02751"  
/translation="KSKQAQAMVQCPSPVAVSQSKPGCYDNGKHVQINQWERTYLG  
NYLCTCGSGRGFNCESEKPEABETCFDKYTGYRVDGTYERDKDSMTWDCTCIGAG  
RGRISCTTAENCHEGGSQYKIGDWRRPHTGTGVMLECVCLNGKGEMTCCKIAEFCF  
DHAACTSVVGSTWBKPQGMWMDVDTCLGESGRITCTSRNKDNODRTSRIVGIDT  
WSKNRNGLQCICGTNGRGECBKERITSUPTSSGGSPFTDVRAAVQPPOPHPOP  
PYGHCTUSGVYVCMGNLKQCNKMCLCTCLGVNSQCEATVCTYCGNLNGSECV  
LPFTYNGRTFYSCYTEGGDHLGMLTSNTSYEQDQIFECTDHYLVLTQGNLSNAL  
CHFFPLYNHNNDTCTSEGRDNMCWGTCTAYADKFQCFPMAAHEEICITNEGVM  
YRIQPDWKHDHMHWRCCTVGNRGSGWTCTAYSQLRDQCIVDDITYNVNDTFKRH  
BEHMLNCQTFCQGRGWKCDDPVDCQSDETCITFYIGDSWBKYHGVRVYCQYGRG  
IEGMHCILQTVPSSSGPVFVEITBTSPQSNHPIQMNAPQPSHSKYLRWRPKNSV  
GEWKEATIPHLNASTIKLGPVVYGOLISIQQYGHQEVTRFDFTTTSTPTPEVS  
TVTGSTTPSPLVASSESVEITASFSVGSWSVASDTVSGSFARVESLESEGDEPOLV  
LPSTGTPSIDLLPGRKIYVNIQYSDVSGSELNLTNPANSVLUSDPPQGVQVNIITAY  
SIWVRSPQAPITGYRISPSVSGSSTELNLTNPANSVLUSDPPQGVQVNIITAY  
BENQSTVPVQQETITGPRDSPRDLQFVTEDVKVTLIMTPPEPAAGYRVYD  
IIPVNLGSHGQLPI.SRNTFAEVTGLSPGTVYFKVFAVSHGRESKPLTAQOTKLDA  
PTNLQFVNETSRLP.PRAAQITVGLTRARGOPQVNVGPSYSKYPLRLNQ  
PASVYTSIVAIFKNGESPKAQVFTLQPGSSIIPYNTETETIVITWTPAIRJVP  
KLGVAPSGGEPANRVEDSGSIVVUGLTPGVYYITQVLRDQGERDAPLVNKVYTP  
LSPPNTHLEANPDVGLVITMSERSTTDITGRIITFTINGQQGNSLEEVHADDOSS  
CTFDNLPLEYNSVYTKDDKESVP.LSDTIIPAVPPTDURLFTNIGDTPMRVWTAP  
PPSIDLNTFLVYSVQNEEDVAELSIPSNNAVLTNLLPGTEYVVSVSVEEHES  
TPLRQKTGLDSPDFSDITANSFVHTAPATITGYRIHRHFSGRPREDR  
VPHRSNTLTLNLTPTGETSVISALNGREESPLLIGQOSTVDVPRDLVEVAATPTS  
ILLI.SDPA.PAVTVRYIRYGETGNGSPVQEFTVPGSKSTATISGLKPGVDYITVTVY  
NGFGPTKTAPGDPTMEIGLQFVETVSVYAQNPGESQPLVQTAVTNIDREKQ  
LAFTDVDVSLDIKAWESPQGVRVYRYSPEDEDGHELFPAPDGEEETAELQGLRPK  
SEYTVVVALDHMMESQLICTQATPAIPTDRLKETQVTFPSTLSAQTNPVLTQLYG

```

VRVPEKGTGMEKINLAPDSSVWVSGLMVATKYEVSVVVKDITLTSRPAQGVVTLT
ENVSPPRAVTDATETIIFISWRTETITGPOVDAVPANGOTPIORTIKEDVRSYT
ITGLOPCTDYKLYLNDNARSPPVUIDASTAIDAPSNLRLATTPNSLLVWOPPR
ARITGIIKXKELGSGSPREVPRPGVTEATITGLEPTEXTIYVIALKNQKSEPL
IGRKTDELQOLVTLPHNPEIILDVSTVOKTFVTHPGIDTNGIQLFQTSQQ
PSVQQMIFEEHRTPTPTTATPIRHRPRPPVNGQEAUSQTTISWAPFQDSEY
IISHPVGTDEEPLRPTPTSAITLGLTRGATYNIIVEALKDQORHKVREVVTV
GNSVNEGLNPTDSDPDPYTVSHYAVGDEWMSKFLKQCLGLGSGHFRCDSS
RWCHDNGVYKIGEKVDRQENQWMSCTCLGNKGFEKCDPHEATCYDDGKTYHVE
QWQEXLGAICSTCTCGQGRGWCDCNRRPGGSPSEGITGGSYQVSRVQRTNTN
VNCPIECMPDLVDQADREDSRE"
14. .6985
/product="fibronectin precursor"
/notes="(pot. aa 1-2324)"
74. .736
/notes="region of internal homology I (5 subunits); domain
with fibrin-, heparin-, s.aureus-binding activity"
842. 952
/notes="region of internal homology I (1 subunit); domain
with collagen-binding activity"
/notes="region of internal homology II (2 subunits) domain
with collagen-binding activity"
1328. 1744
/notes="region of internal homology I (3 subunits) domain
with collagen-binding activity"
1745. .2020
/notes="region of internal homology III (1 subunit)"
2075. .5893
/notes="region of internal homology III (14 subunits)"
2347. .3175
/notes="domain with DNA-binding activity"
4490. 4501
/notes="cell binding site"
4811. .5080
/notes="ED region for alternative splicing"
5081. .5893
/notes="domain with heparin-binding activity"
5894. .6160
/notes="III CS region for alternative splicing"
6161. .6373
/notes="region of internal homology III (1 subunit)"
6442. .6838
/notes="region of internal homology I (3 subunits); domain
with fibrin-binding activity"
7679
polyA_site
ORIGIN
Query Match 95.2%; Score 7423; DB 9; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 349 GAAGAGCAAGGAGGAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCACTCA 408
DB 1 GAAAGAGCAAGAGGAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCACTCA 60
QY 409 AAGCAAGCCCGGTTGTTATGACAAATGGAATAACATCATCAGATAAATCAACAGTGGGAGCG 468
DB 61 AAGCAAGCCCGGTTGTTATGACAAATGGAATAACATCATCAGATAAATCAACAGTGGGAGCG 120
QY 469 GACCTACCTAGCAATGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 528
DB 121 GACCTACCTAGCAATGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
QY 529 CGAAAGTAACTGAAGCTGAAGAGACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 588
DB 181 CGAAAGTAACTGAAGCTGAAGAGACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
QY 589 AGTGGGTGACATTATGACGCTTAAAGACTCCAGACTGTTGGAGCTGTACCTCATCGG 648
DB 241 AGTGGGTGACATTATGACGCTTAAAGACTCCAGACTGTTGGAGCTGTACCTCATCGG 300
QY 649 GGCTGGGAGGAGGAGATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAATC 708

```

```

301 GGCTGGGAGGAGGAGATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGTCACTC 360
QY 709 CTACAAGATTGTTGACACCTCGAGGAGACCAATGAGACTGGTGTGTACATGTTTAGAGTG 768
DB 361 CTACAAGATTGTTGACACCTCGAGGAGACCAATGAGACTGGTGTGTACATGTTTAGAGTG 420
QY 769 TGTGTGTTCTTGGTAATGGAAGAGAGAAATGACCTGCAAGCCCATAGCTGAGAAAGTGT 828
DB 421 TGTGTGTTCTTGGTAATGGAAGAGAGAAATGACCTGCAAGCCCATAGCTGAGAAAGTGT 480
QY 829 TGATCATGCTGCTGGGACTTCTTATGTTGTCGAGAAACGTTGGGAGAAAGCCCTACCAAGG 888
DB 481 TGATCATGCTGCTGGGACTTCTTATGTTGTCGAGAAACGTTGGGAGAAAGCCCTACCAAGG 540
QY 889 CTGGATGATGTTAGATTGTTACTTTCCTGGGAGAGGAGGAGCGGACGCTACCTTGGACTTC 948
DB 541 CTGGATGATGTTAGATTGTTACTTTCCTGGGAGAGGAGGAGCGGACGCTACCTTGGACTTC 600
QY 949 TAGAATAGATGCAACGATCAGGACACAGGACACATCTATAGAAATTTGGAGACACTGGAG 1008
DB 601 TAGAATAGATGCAACGATCAGGACACAGGACACATCTATAGAAATTTGGAGACACTGGAG 660
QY 1009 CAAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCAAGGCAACGCCCGGAGAGA 1068
DB 661 CAAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCAAGGCAACGCCCGGAGAGA 720
QY 1069 GTGGAGTGTGAGAGGACACATCTGTCGACACACACATCGAGCGGATCTGGCCCTTCAC 1128
DB 721 GTGGAGTGTGAGAGGACACATCTGTCGACACACACATCGAGCGGATCTGGCCCTTCAC 780
QY 1129 CGATGTTGCTGAGCTGTTTACCAACCGCAGCTCACCACCGAGCTCTCCCTATGGCCA 1188
DB 781 CGATGTTGCTGAGCTGTTTACCAACCGCAGCTCACCACCGAGCTCTCCCTATGGCCA 840
QY 1189 CTGTGTACAGACAGTGTGTGGTCTACTCTGTGGGATGAGTGGCTGAGAGACACAGG 1248
DB 841 CTGTGTACAGACAGTGTGTGGTCTACTCTGTGGGATGAGTGGCTGAGAGACACAGG 900
QY 1249 AAATAAGCAAAATGCTTTCAGCTGCTGGGCAACCGAGTCACTGCCAAGAGACAGCTGT 1308
DB 901 AAATAAGCAAAATGCTTTCAGCTGCTGGGCAACCGAGTCACTGCCAAGAGACAGCTGT 960
QY 1309 AACCCAGACTTACGTTGGCAACTCAATGAGAGAGCCATGTGTCTTACCATTCACCTACAA 1368
DB 961 AACCCAGACTTACGTTGGCAACTTAAATGAGAGCCATGTGTCTTACCATTCACCTACAA 1020
QY 1369 TGCGAGGAGCTTCTACTCTCTGCACACAGAGGGGCGACAGGACGACATCTTGGTGCAG 1428
DB 1021 TGCGAGGAGCTTCTACTCTCTGCACACAGAGGGGCGACAGGACGACATCTTGGTGCAG 1080
QY 1429 CACAACTTCGAAATTAAGACAGGACCAAGAAATACCTCTTCTGCAAGACCACTGTGTTT 1488
DB 1081 CACAACTTCGAAATTAAGACAGGACCAAGAAATACCTCTTCTGCAAGACCACTGTGTTT 1140
QY 1489 GGTTCAGACTCCAGAGGAGAAATTCGAATGGTGGCTGTGGCCACTTCCCTCTTATACAA 1548
DB 1141 GGTTCAGACTCCAGAGGAGAAATTCGAATGGTGGCTGTGGCCACTTCCCTCTTATACAA 1200
QY 1549 CAAACCAAAATTAACCTGATTGCACTTCTGAGGCGAGAGAGACAAATGAAGTGTGTGG 1608
DB 1201 CAAACCAAAATTAACCTGATTGCACTTCTGAGGCGAGAGAGACAAATGAAGTGTGTGG 1260
QY 1609 GACCAACAGAACTATGATGCGGACCAAGATTTGGGTTCTGCCCATGGCTGGCCACGA 1668
DB 1261 GACCAACAGAACTATGATGCGGACCAAGATTTGGGTTCTGCCCATGGCTGGCCACGA 1320
QY 1669 GGAAATCTGCACAAACCAATGAAGGGGTGATGTACCGCATTTGGAGATCAGTGGGATGA 1728
DB 1321 GGAAATCTGCACAAACCAATGAAGGGGTGATGTACCGCATTTGGAGATCAGTGGGATGA 1380
QY 1729 GCATGACATGGGTGATGATGAGGTGACCTGTGTGGGATGCTCGTGGGGATGAGC 1788

```

1381	GCATGACATGGGTCCACATGATGAGGTGCGACGTGTGTTGGGAATATGGTGTGGGGGAATGGAC	1440
1789	ATGCATTGCCTACTCCGACGTTCCAGATCAGTGCATTGTTGATGACATCACTTACAATGT	1848
1441	ATGCATTGCCTACTCGCAACTTCGAGATCAGTGCATTGTTGATGACATCACTTACAATGT	1500
1849	GAACGACACATTTCCACAAGCGTTCATCAAGAGGGGCACATGCTGNACTGTACATGCTTCGG	1908
1501	GAACGACACATTTCCACAAGCGTTCATCAAGAGGGGCACATGCTGNACTGTACATGCTTCGG	1560
1909	TCAGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAAATGCCAGGATTCAGAGACTGG	1968
1561	TCAGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAAATGCCAGGATTCAGAGACTGG	1620
1969	GAGCTTTTATCAAAATGGAGATTCAATGGGAGAGTATGTGCATGGTGTGCAGATACCAGTG	2028
1621	GAGCTTTTATCAAAATGGAGATTCAATGGGAGAGTATGTGCATGGTGTGCAGATACCAGTG	1680
2029	CTACTGCTATGGCCGTGGCANTGGGAGTGGCANATGCCAACCTTTTACAGACCTATCCAAAG	2088
1681	CTACTGCTATGGCCGTGGCANTGGGAGTGGCANATGCCAACCTTTTACAGACCTATCCAAAG	1740
2089	CTCAAGTGGTCCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCAGGCCCAACTCCCACCC	2148
1741	CTCAAGTGGTCCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCAGGCCCAACTCCCACCC	1800
2149	CATCCAGTGGAAATGCAACCAAGCCATCTCACATTTCCAAAGTACATTTCCAGTGGAGACC	2208
1801	CATCCAGTGGAAATGCAACCAAGCCATCTCACATTTCCAAAGTACATTTCCAGTGGAGACC	1860
2209	TAAAAATTCGTAGGCCGTGGGAAGGAAGCTACCATACCAAGGCCACTTAAATCTCCTTACAC	2268
1861	TAAAAATTCGTAGGCCGTGGGAAGGAAGCTACCATACCAAGGCCACTTAAATCTCCTTACAC	1920
2269	CATCAAAGGCTGAAGCCTGGTGTGTTATAGAGGGCCAGCTCATCAGCATCCAGCAGTA	2328
1921	CATCAAAGGCTGAAGCCTGGTGTGTTATAGAGGGCCAGCTCATCAGCATCCAGCAGTA	1980
2329	CGGCCACCAAGAAAGTCACTCGCTTTGACTTTCACTTCAACCAACAGCAGCAGCACCTGTGAC	2388
1981	CGGCCACCAAGAAAGTCACTCGCTTTGACTTTCACTTCAACCAACAGCAGCAGCACCTGTGAC	2040
2389	CAGCAACACGGTGACAGGAGAGACGACTCCCTTTTCTCTTTGTGGCCACTTCTGAAATC	2448
2041	CAGCAACACGGTGACAGGAGAGACGACTCCCTTTTCTCTTTGTGGCCACTTCTGAAATC	2100
2449	TGTGACCGGAAATCACAGCCAGTAGCTTGTGGTCTCTCGGCTCTCAGCTTCGGACACCGT	2508
2101	TGTGACCGGAAATCACAGCCAGTAGCTTGTGGTCTCTCGGCTCTCAGCTTCGGACACCGT	2160
2509	GTCCGATTTCCGGGTGGAAATATGACTGATGAGGAGGGAGATGAGCCACAGTACTGGA	2568
2161	GTCCGATTTCCGGGTGGAAATATGACTGATGAGGAGGGAGATGAGCCACAGTACTGGA	2220
2569	TCTTTCCAAGCACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTCTGGCCGAAATATACAT	2628
2221	TCTTTCCAAGCACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTCTGGCCGAAATATACAT	2280
2629	TGTAATATGTTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCCTGTCTACTTTACA	2688
2281	TGTAATATGTTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCCTGTCTACTTTACA	2340
2689	AACAAACAGCCCTGATGCCCCCTCCTGACCCGAATGTGGACCAAGTTTGATGACACCTCAAT	2748
2341	AACAAACAGCCCTGATGCCCCCTCCTGACCCGAATGTGGACCAAGTTTGATGACACCTCAAT	2400
2749	TGTTGTTCCGTGGAGACAGCCAGGCTCCATCAAGGGTACAGAAATAGTCTATTTCGCC	2808
2401	TGTTGTTCCGTGGAGACAGCCAGGCTCCATCAAGGGTACAGAAATAGTCTATTTCGCC	2460
2809	ATCAGTAGAAGGTAGCAGCACAGAAGCTCAACCTTCTGAAACTGCAAACTCCGTCACCT	2868
2461	ATCAGTAGAAGGTAGCAGCACAGAAGCTCAACCTTCTGAAACTGCAAACTCCGTCACCT	2520

[illegible]







Db 5761 GGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTATTACTGGCTGGAACCGGGAAC 5820  
 2y 6169 CGAATATACAAATTTATGTCTATGCGCTGGAAGATAATATCAGAGAGGAGCCCTGATGTTG 6228  
 Db 5821 CGAATATACAAATTTATGTCTATGCGCTGGAAGATAATATCAGAGAGGAGCCCTGATGTTG 5880  
 2y 6229 AAGGAAAAAGACAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCATGG 6288  
 3b 5881 AAGGAAAAAGACAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCATGG 5940  
 2y 6289 ACCAGAGATCTTGATGTTCTCTCCACAGTTCAAAAGACCCCTTTCGTCAACCACCCCTGG 6348  
 3b 5941 ACCAGAGATCTTGATGTTCTCTCCACAGTTCAAAAGACCCCTTTCGTCAACCACCCCTGG 6000  
 2y 6349 GTATGACACTGGAAATGGTATTCAGCTTCTGGCACTTCTGGTACAGAACCCCAAGTGTGG 6408  
 3b 6001 GTATGACACTGGAAATGGTATTCAGCTTCTGGCACTTCTGGTACAGAACCCCAAGTGTGG 6060  
 2y 6409 GCAACAAATGATCTTTGAGGAACATGTTTTAGCGGACACACCCGCCCAACACCGCCAC 6468  
 Db 6061 GCAACAAATGATCTTTGAGGAACATGTTTTAGCGGACACACCCGCCCAACACCGCCAC 6120  
 2y 6469 CCCCATAGGCAATAGGCAAGACATACCGCGAATGTAGGACAGAGCTCTCTCTCA 6528  
 Db 6121 CCCCATAGGCAATAGGCAAGACATACCGCGAATGTAGGACAGAGCTCTCTCTCA 6180  
 2y 6529 GACAAACATCTCATGGGCCCCCATTTCCAGGACACTTCTGAGTACATCATTTTCATGTCTATCC 6588  
 Db 6181 GACAAACATCTCATGGGCCCCCATTTCCAGGACACTTCTGAGTACATCATTTTCATGTCTATCC 6240  
 2y 6589 TGTGGCACTGATGAAGAACCTTTACAGTTTCAGGGTTCCTGGAACTTCTACAGTGGCCAC 6648  
 Db 6241 TGTGGCACTGATGAAGAACCTTTACAGTTTCAGGGTTCCTGGAACTTCTACAGTGGCCAC 6300  
 2y 6649 TCTGACAGGCTCACCAGAGGTGCCACTCAACATCATAGTGGAGGCACTGAGAAACCCA 6708  
 Db 6301 TCTGACAGGCTCACCAGAGGTGCCACTCAACATCATAGTGGAGGCACTGAGAAACCCA 6360  
 2y 6709 GCAGAGGATTAAGTTTCGGGAAGAGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6768  
 Db 6361 GCAGAGGATTAAGTTTCGGGAAGAGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6420  
 2y 6769 GAACCAACTACGATGACTCGTCTTTGACCCCTACACAGTGTCCCATATATCCCGTTGG 6828  
 Db 6421 GAACCAACTACGATGACTCGTCTTTGACCCCTACACAGTGTCCCATATATCCCGTTGG 6480  
 2y 6829 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAACTGTGTGGCAGTGTAGGCTT 6888  
 Db 6481 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAACTGTGTGGCAGTGTAGGCTT 6540  
 2y 6889 TGGAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGGCATCACAATGGTGTGAACCTA 6948  
 Db 6541 TGGAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGGCATCACAATGGTGTGAACCTA 6600  
 2y 6949 CAAGATTGGAGAGAGTGGGACCGTCCAGGAGAAATGGCCAGATGATGAGCTGCACATG 7008  
 Db 6601 CAAGATTGGAGAGAGTGGGACCGTCCAGGAGAAATGGCCAGATGATGAGCTGCACATG 6660  
 2y 7009 TCTTGGGACCGGAAAGGAGAAATTCAGATGTGACCTCATGAGGCAACGTTTACGATGA 7068  
 Db 6661 TCTTGGGACCGGAAAGGAGAAATTCAGATGTGACCTCATGAGGCAACGTTTACGATGA 6720  
 2y 7069 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTTGGCTC 7128  
 Db 6721 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTTGGCTC 6780  
 2y 7129 CTGCACATGCTTTGGAGCCAGCGGGCTGGCGTGTGACAACTGGCCGACAGCTGGGGG 7188  
 Db 6781 CTGCACATGCTTTGGAGCCAGCGGGCTGGCGTGTGACAACTGGCCGACAGCTGGGGG 6840  
 2y 7189 TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTACAAACAGTATTTCTCAGAGATACCA 7248  
 Db 6841 TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTACAAACAGTATTTCTCAGAGATACCA 6900

QY 7249 TCAGAGAACAAACACTAATGTTAAATTTGCCAAATTTGAGTGCTTCATGCCCTTTAGATGTACA 7308  
 Db 6901 TCAGAGAACAAACACTAATGTTAAATTTGCCAAATTTGAGTGCTTCATGCCCTTTAGATGTACA 6960  
 QY 7309 GGCTCACAGAGAAGATTCCTCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGT 7368  
 Db 6961 GGCTCACAGAGAAGATTCCTCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGT 7020  
 QY 7369 CTCTCTGCAAGATCCATCTCTAACTGGAGTGAATGTAGCAGACCCAGCTTAGAGTTCTTC 7428  
 Db 7021 CTCTCTGCAAGATCCATCTCTAACTGGAGTGAATGTAGCAGACCCAGCTTAGAGTTCTTC 7080  
 QY 7429 TTTCTTTCTTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT 7488  
 Db 7081 TTTCTTTCTTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT 7140  
 QY 7489 TCTCAAGCATCACCCCTGGAGTTTCTCGAGGTTTCTCATAAATGAGGCTGCACATT 7548  
 Db 7141 TCTCAAGCATCACCCCTGGAGTTTCTCGAGGTTTCTCATAAATGAGGCTGCACATT 7200  
 QY 7549 GCCTGTTCTGCTTCCAGTATTTCAATACCGCTCAGTATTTTAAATGAGTGAATTTCTAAGA 7608  
 Db 7201 GCCTGTTCTGCTTCCAGTATTTCAATACCGCTCAGTATTTTAAATGAGTGAATTTCTAAGA 7260  
 QY 7609 TTTGTTTGGGATCAATAGGAAGCATATGAGGCAACCAAGATGCAAAATGTTTGAAT 7668  
 Db 7261 TTTGTTTGGGATCAATAGGAAGCATATGAGGCAACCAAGATGCAAAATGTTTGAAT 7320  
 QY 7669 GATATGACCAAAATTTTAAAGTAGGAAGTACCANAACACTCTGCTTCACTTAAAGTGT 7728  
 Db 7321 GATATGACCAAAATTTTAAAGTAGGAAGTACCANAACACTCTGCTTCACTTAAAGTGT 7380  
 QY 7729 CTGGCCCCCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7788  
 Db 7381 CTGGCCCCCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7440  
 QY 7789 CAGTACT 7795  
 Db 7441 CAGTACT 7447

## RESULT 13

LOCUS A14133 7705 bp DNA linear PAT 30-NOV-1994  
 DEFINITION Fibronectin gene.  
 ACCESSION A14133  
 VERSION A14133.1 GI:490127  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 7705)  
 AUTHORS Baralle, F.E.  
 TITLE Fibronectins  
 JOURNAL Patent: EP 0207751-A 2 07-JAN-1987;  
 Delta Biotechnology Limited  
 FEATURES  
 source  
 1..7705  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 misc\_feature  
 5..6985  
 /note="CDS, fibronectin gene"  
 misc\_feature  
 647..649  
 /note="the sequence given is TGA, but the codon is shown  
 in the specification as encoding Glycine for which the  
 correct codon is GGN"

## ORIGIN

Query Match 95.2%; Score 7418.2; DB 6; Length 7705;  
 Best Local Similarity 99.8%; Pred. No. 0;

Matches 7429; Conservative 0; Mismatches 18; Indels 0; Gaps 0;			
QY	349	GAAGAGCAAGAGGCAAGGCTCAGCAAAATGGTTGAGCCCGGTCAGTCA	408
Db	1	GAAGAGCAAGAGGCAAGGCTCAGCAAAATGGTTGAGCCCGGTCAGTCA	60
QY	409	AAGCAAGCCGGTCTGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCG	468
Db	61	AAGCAAGCCGGTCTGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCG	120
QY	469	GACCTACCTAGCAAAAGCGTTGGTTGTTACTGTTATCGAGGAAGCCGAGGTTTAACTG	528
Db	121	GACCTACCTAGGTAATGTTGGTTGTTACTGTTATCGAGGAAGCCGAGGTTTAACTG	180
QY	529	CGAAGTAAACCTGAAGCTGAAGACACTGCTTTGACAAGTACACTGGGAACACTTACCG	588
Db	181	CGAAGTAAACCTGAAGCTGAAGACACTGCTTTGACAAGTACACTGGGAACACTTACCG	240
QY	589	AGTGGGTGACACTTATGAGCGCTCTTAAAGACTCCATGATCTGGGACTGTACTGCAATCGG	648
Db	241	AGTGGGTGACACTTATGAGCGCTCTTAAAGACTCCATGATCTGGGACTGTACTGCAATCGG	300
QY	649	GCCTGGGCGAGGAGAAATGACTGTACCATCGCAACCGCTGCCATGAAGGGGTCACTC	708
Db	301	GCCTGGGCGAGGAGAAATGACTGTACCATCGCAACCGCTGCCATGAAGGGGTCACTC	360
QY	709	CTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	768
Db	361	CTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	420
QY	769	TGTGTGCTTTGGTAATGGAAAGAGAAATGAGACTGTGAGCCCATGAGTGAAGAGTGT	828
Db	421	TGTGTGCTTTGGTAATGGAAAGAGAAATGAGACTGTGAGCCCATGAGTGAAGAGTGT	480
QY	829	TGATCATGCTGGTGGACTCTCTATGTTGGTCCGAGAAACGTTGGGAGAACCTTACCAAGG	888
Db	481	TGATCATGCTGGTGGACTCTCTATGTTGGTCCGAGAAACGTTGGGAGAACCTTACCAAGG	540
QY	889	CTGGATGATGTTAGTGTACTGCTGGGAGAGAGGAGCGGACGTCATCTGCACTTC	948
Db	541	CTGGATGATGTTAGTGTACTGCTGGGAGAGAGGAGCGGACGTCATCTGCACTTC	600
QY	949	TAGAATAGATGCAAGATCAGGACACAGGACATCTTATAGAATTTGAGACACCTGGAG	1008
Db	601	TAGAATAGATGCAAGATCAGGACACAGGACATCTTATAGAATTTGAGACACCTGGAG	560
QY	1009	CAAGAAGATAATCGAGGAAAACCTGCTCAGTGCATCTGCA CAGGCAACCGGCGAGGAGA	1068
Db	661	CAAGAAGATAATCGAGGAAAACCTGCTCAGTGCATCTGCA CAGGCAACCGGCGAGGAGA	720
QY	1069	GTGGAGTGTGAGAGGACACCTCTGTGCGAGACCATCTGAGCGGATCTGGCCCTTCAC	1128
Db	721	GTGGAGTGTGAGAGGACACCTCTGTGCGAGACCATCTGAGCGGATCTGGCCCTTCAC	780
QY	1129	CGATGTTCTGTCAGCTGTTTACCAACCGCAGGCTCACCCCCAGCTCTCCCTATGGCCA	1188
Db	781	CGATGTTCTGTCAGCTGTTTACCAACCGCAGGCTCACCCCCAGCTCTCCCTATGGCCA	840
QY	1189	CTGTGTCA CAGACAGTGGTGTGTTCTCTGTGGGATGCACTGGCTGAAGACACAGG	1248
Db	841	CTGTGTCA CAGACAGTGGTGTGTTCTCTGTGGGATGCACTGGCTGAAGACACAGG	900
QY	1249	AAATAAGCAAAATGCTTTGACGCTGCTGGGCAACCGAGTCAGTCCGAGAGACAGCTGT	1308
Db	901	AAATAAGCAAAATGCTTTGACGCTGCTGGGCAACCGAGTCAGTCCGAGAGACAGCTGT	960
QY	1309	AACCCAGACTTACGTTGGGCACTCAAAATGGAGAGCCATGTCTTACCATTCACCTACAA	1368
Db	961	AACCCAGACTTACGTTGGGCACTCAAAATGGAGAGCCATGTCTTACCATTCACCTACAA	1020
QY	1369	TGGCAGGACGCTTCTACTCTCTGACCA CAGAAAGGGCGACAGGACGATCTTTGGTGCAG	1428
Db	1021	TGGCAGGACGCTTCTACTCTCTGACCA CAGAAAGGGCGACAGGACGATCTTTGGTGCAG	1080

QY	1429	CACAACTTCGAATTTATGAGAGGACCAAGAAATACCTTTCTGCACAGACCACTGTTTT	1488
Db	1081	CACAACTTCGAATTTATGAGAGGACCAAGAAATACCTTTCTGCACAGACCACTGTTTT	1140
QY	1489	GGTTCAGACTCGAGGAGGAAATTCCAATGGTCCCTGTGTGCCACTTCCCTTCTATACAA	1548
Db	1141	GGTTCAGACTCAAGGAGGAAATTCCAATGGTCCCTGTGTGCCACTTCCCTTCTATACAA	1200
QY	1549	CAACCAAAATTCACCTGATTGCACTTCTGAGGGCAGAGAGACAAATCAAGTGGTGTGG	1608
Db	1201	CAACCAAAATTCACCTGATTGCACTTCTGAGGGCAGAGAGACAAATCAAGTGGTGTGG	1260
QY	1609	GACCAACAGAACTATGATGCCGAGCCAGAAAGTTTGGGTTCTGCCCATGGCTGCCCAAG	1668
Db	1261	GACCAACAGAACTATGATGCCGAGCCAGAAAGTTTGGGTTCTGCCCATGGCTGCCCAAG	1320
QY	1669	GGAAATCTGCACAAACCAATGAAGGGGTCACTGACCGCATTTGGAGATCACTGGGATAAGCA	1728
Db	1321	GGAAATCTGCACAAACCAATGAAGGGGTCACTGACCGCATTTGGAGATCACTGGGATAAGCA	1380
QY	1729	GCATGACATGGGTCACTGATGAGGTGCACTGTGTGGGAAATGGTCTGGGGAATGGAC	1788
Db	1381	GCATGACATGGGTCACTGATGAGGTGCACTGTGTGGGAAATGGTCTGGGGAATGGAC	1440
QY	1789	ATGCATTTGCTTACTCCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTACATGT	1848
Db	1441	ATGCATTTGCTTACTCCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTACATGT	1500
QY	1849	GAAACGACACATTCACAAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG	1908
Db	1501	GAAACGACACATTCACAAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTCGG	1560
QY	1909	TCAGGTTCGGGGCAGGTGGAAAGTGTATCCCGTCGACCAATGCCAGGATTCAGAGACTGG	1968
Db	1561	TCAGGTTCGGGGCAGGTGGAAAGTGTATCCCGTCGACCAATGCCAGGATTCAGAGACTGG	1620
QY	1969	GAGCTTTTATCAAAATGGAGATTTCATGGAGAAAGTATGTCATGGTGTGATACATCCAGTG	2028
Db	1621	GAGCTTTTATCAAAATGGAGATTTCATGGAGAAAGTATGTCATGGTGTGATACATCCAGTG	1680
QY	2029	CTACTGCTATGGCCGTGGCAATTTGGGAGTGGCAATTCGCCAAACCTTTACAGACCTATCCAAG	2088
Db	1681	CTACTGCTATGGCCGTGGCAATTTGGGAGTGGCAATTCGCCAAACCTTTACAGACCTATCCAAG	1740
QY	2089	CTCAAGTGGTCTCTGAAAGTATTTATCACTGAGACTCCGAGTCCAGCCCAACTCCCAACC	2148
Db	1741	CTCAAGTGGTCTCTGAAAGTATTTATCACTGAGACTCCGAGTCCAGCCCAACTCCCAACC	1800
QY	2149	CATCCAGTGGAAATGCACCAAGCCATCTCAATTTCCAAAGTACATTTCTCAGGTGGAGACC	2208
Db	1801	CATCCAGTGGAAATGCACCAAGCCATCTCAATTTCCAAAGTACATTTCTCAGGTGGAGACC	1860
QY	2209	TAAAAATTTCTAGGCGGTGGAGAGAGCTACCATCCAGGCCACTTAAACTCTCTACAC	2268
Db	1861	TAAAAATTTCTAGGCGGTGGAGAGAGCTACCATCCAGGCCACTTAAACTCTCTACAC	1920
QY	2269	CATCAAAAGGCTTGAAGCCTGTGTGTATACGAGGCGCAGCTCATCAGCATCCAGCAGTA	2328
Db	1921	CATCAAAAGGCTTGAAGCCTGTGTGTATACGAGGCGCAGCTCATCAGCATCCAGCAGTA	1980
QY	2329	CGGCCACCAAGAGTGAATGCTGCTTGAATTCACCAACAGCAGCAGCAGCAGCAGTGC	2388
Db	1981	CGGCCACCAAGAGTGAATGCTGCTTGAATTCACCAACAGCAGCAGCAGCAGCAGTGC	2040
QY	2389	CAGCAACACCGTGCACAGGAGAGCGACTCCCTTTCTCTTTGGGCCACTTCTGAATC	2448
Db	2041	CAGCAACACCGTGCACAGGAGAGCGACTCCCTTTCTCTTTGGGCCACTTCTGAATC	2100
QY	2449	TGTGAACGAAATCAGACCGAGTGTGTGTGTCTCTCTGGGTCTCAGCTTCCGACACCGT	2508
Db	2101	TGTGAACGAAATCAGACCGAGTGTGTGTGTCTCTCTGGGTCTCAGCTTCCGACACCGT	2160

2509 GTCCGATTCCGGTGGAAATATAGCTGAGTGGAGGAGAGATGAGCCACAGTACTGGA 2568  
2161 GTCCGATTCCGGTGGAAATATAGCTGAGTGGAGGAGAGATGAGCCACAGTACTGGA 2220  
2569 TCTTCCAGACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAATACAT 2628  
2221 TCTTCCAGACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAATACAT 2280  
2629 TGTAAATGCTATCAGATATCTCAGATGAGGAGAGAGATTTGATCTCTGTACTTACA 2688  
2281 TGTAAATGCTATCAGATATCTCAGATGAGGAGAGAGATTTGATCTCTGTACTTACA 2340  
2689 AACACAGCCCTGATGCTCTGACCGAGCTGTGGACCAAGTTGATGACACCTCAAT 2748  
2341 AACACAGCCCTGATGCTCTGACCGAGCTGTGGACCAAGTTGATGACACCTCAAT 2400  
2749 TGTGTCTGCTGAGCAGACCCAGGCTCCATCAGAGGTACAGATAGTCTNTTCGCC 2808  
2401 TGTGTCTGCTGAGCAGACCCAGGCTCCATCAGAGGTACAGATAGTCTNTTCGCC 2460  
2809 ATCAGTAGAAGGTAGCAGCAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCT 2868  
2461 ATCAGTAGAAGGTAGCAGCAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCT 2520  
2869 CAGTGACTTGCACCTGCTTTCAGTATACATCACTATCTATGCTGTGAGAGAAATCA 2928  
2521 CAGTGACTTGCACCTGCTTTCAGTATACATCACTATCTATGCTGTGAGAGAAATCA 2580  
2929 AGAAGTACACCTGCTTCAATCAAGAAACCACTGGCACCCAGCTCAGATACAGT 2988  
2581 AGAAGTACACCTGCTTCAATCAAGAAACCACTGGCACCCAGCTCAGATACAGT 2640  
2989 GCCTCTCCAGGAGCTGAGTTGTGAGTGTGAGAGTGAAGAGTCAAGGTCAATATGTG 3048  
2641 GCCTCTCCAGGAGCTGAGTTGTGAGTGTGAGAGTGAAGAGTCAAGGTCAATATGTG 2700  
3049 GACACCGCTGAGAGTGCAGTGCAGCTGACCGCTACCTGTGAGTGTGAGTGTGAGTGC 3108  
2701 GACACCGCTGAGAGTGCAGTGCAGCTGACCGCTACCTGTGAGTGTGAGTGTGAGTGC 2760  
3109 TGGGAGACGGGAGAGGCTGCCATCAGAGAAACACTTTGAGAGAGTCAAGGTCAAGGT 3168  
2761 TGGGAGACGGGAGAGGCTGCCATCAGAGAAACACTTTGAGAGAGTCAAGGTCAAGGT 2820  
3169 GTCCCTGGGCTCACTATTAATTTCAAGTCTTTGAGTGCAGTGCAGTGCAGTGCAGTGC 3228  
2821 GTCCCTGGGCTCACTATTAATTTCAAGTCTTTGAGTGCAGTGCAGTGCAGTGCAGTGC 2880  
3229 GCCTCTGACTGTCAACAGACAAACCAACTGGATGCTCCCACTAACCTCCAGTTGTCAA 3288  
2881 GCCTCTGACTGTCAACAGACAAACCAACTGGATGCTCCCACTAACCTCCAGTTGTCAA 2940  
3289 TGAATGATTTACTGCTGTGATGAGTGGATCCACCTCGGCGCCAGATTAACAGGATA 3348  
2941 TGAATGATTTACTGCTGTGATGAGTGGATCCACCTCGGCGCCAGATTAACAGGATA 3000  
3349 CCGACTGACCGTGGGCTTTACCCGAGAGACAGCCAGCCAGTACAAATGTGGGTCCTC 3408  
3001 CCGACTGACCGTGGGCTTTACCCGAGAGAGCCAGCCAGTACAAATGTGGGTCCTC 3060  
3409 TGTCTCAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGCT 3468  
3061 TGTCTTCAAGTACCCCTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGCT 3120  
3469 GGCATTAAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTTACCACACTGCGAGC 3528  
3121 GGCATTAAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTTACCACACTGCGAGC 3180  
3529 TGGAGCTCTATTCCACTTAAACCCAGAGTGAAGTGAAGACCACTTGTGATCAATG 3588  
3181 TGGAGCTCTATTCCACTTAAACCCAGAGTGAAGTGAAGACCACTTGTGATCAATG 3240  
3589 GACGCTGCTCCAGAAATTTGTTTAACTGGGTGTACGACCAAGCCAGGAGGAGGAGC 3648

3241 GACGCTGCTCCAGAAATTTGTTTAACTGGGTGTACGACCAAGCCAGGAGGAGGC 3300  
3649 ACCACGAGAGTGAATCTCAGACTCAGGAAGCATGTTGTGTGCTGGCTTGTACTCCAGAGT 3708  
3301 ACCACGAGAGTGAATCTCAGACTCAGGAAGCATGTTGTGTGCTGGCTTGTACTCCAGAGT 3360  
3709 AGAATACGCTTACACCATCCAAAGTCTCAGAGATGAGACAGAAAGAGATGCGCCAAATGT 3768  
3361 AGAATACGCTTACACCATCCAAAGTCTCAGAGATGAGACAGAAAGAGATGCGCCAAATGT 3420  
3769 AAAAAGAGTGTGACACATTTGTCTCCACAACTTGTGATCTGGAGGCAAAACCTGGA 3828  
3421 AAAAAGAGTGTGACACATTTGTCTCCACAACTTGTGATCTGGAGGCAAAACCTGGA 3480  
3829 CACTGGAGTGTCTCAGCTTCTCTGGAGAGGAGACCAACCCAGACATTAATCTGTTATAG 3888  
3481 CACTGGAGTGTCTCAGCTTCTCTGGAGAGGAGACCAACCCAGACATTAATCTGTTATAG 3540  
3889 AATTACCAACCCCTTCAAAACCGCCAGCAGGGAATTTCTTTGGAGAAAGTGTGCTCATGC 3948  
3541 AATTACCAACCCCTTCAAAACCGCCAGCAGGGAATTTCTTTGGAGAAAGTGTGCTCATGC 3600  
3949 TGATCAGAGCTCTGCACTTTTGTATACCTGAGTCCCGGCTTGGAGTACATGTCAGTGT 4008  
3601 TGATCAGAGCTCTGCACTTTTGTATACCTGAGTCCCGGCTTGGAGTACATGTCAGTGT 3660  
4009 TTACACTGTCAAGGATGACCAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
3661 TTAACACTGTCAAGGATGACCAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
4069 TCCCTCTCCCACTGACCTGCGATTCAACCAATTTGTCAGACCACTGCTGTCACCTG 4128  
3721 TCCCTCTCCCACTGACCTGCGATTCAACCAATTTGTCAGACCACTGCTGTCACCTG 3780  
4129 GGCTCCACCCCTTCAATTTGATTTAACTTCTCTGCTGCTTACTCACTGCTGAAAAA 4188  
3781 GGCTCCACCCCTTCAATTTGATTTAACTTCTCTGCTGCTTACTCACTGCTGAAAAA 3840  
4189 TGAGAGAGATTTGACAGATGTCATTTCTCTTTCAGACCAATGCTGCTTAACTTAA 4248  
3841 TGAGAGAGATTTGTCAGATGTCATTTCTCTTTCAGACCAATGCTGCTTAACTTAA 3900  
4249 TCTCTGCTGCTTACAGATATGATGAGTGTCTCCAGTGTCTTACGAAACCAATGAGAG 4308  
3901 TCTCTGCTGCTTACAGATATGATGAGTGTCTCCAGTGTCTTACGAAACCAATGAGAG 3960  
4309 CACACTCTTAGAGGAGACAGAAAAACAGTCTTTGATTTCCCACTGGCAATGCTTTTC 4368  
3961 CACACTCTTAGAGGAGACAGAAAAACAGTCTTTGATTTCCCACTGGCAATGCTTTTC 4020  
4369 TGATATTACTGCCAACTCTTTTACTGTGCACTGATTTGCTCTCGAGCCACCATCTGG 4428  
4021 TGATATTACTGCCAACTCTTTTACTGTGCACTGATTTGCTCTCGAGCCACCATCTGG 4080  
4429 CTAAGGATCCGCTCATCCGAGCATCTCAGTGGAGACCTTCGAGAGATCGGGTGC 4488  
4081 CTAAGGATCCGCTCATCCGAGCATCTCAGTGGAGACCTTCGAGAGATCGGGTGC 4140  
4489 CCACTCTCGGAAATTCATCACCTTCAACCTTCACTCCAGGACACAGATGCTGCTCAG 4548  
4141 CCACTCTCGGAAATTCATCACCTTCAACCTTCACTCCAGGACACAGATGCTGCTCAG 4200  
4549 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTTGATTTGGCCAACTCAACAGT 4608  
4201 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTTGATTTGGCCAACTCAACAGT 4260  
4609 TTTCTGATTTCCGAGGACCTTGGAAAGTGTGTCGACCCCTTCACTCCAGGACACAGATGCTGCTCAG 4668  
4261 TTTCTGATTTCCGAGGACCTTGGAAAGTGTGTCGACCCCTTCACTCCAGGACACAGATGCTGCTCAG 4320  
4669 CTGGGATGCTCTCTCTGCTCAGTGTGATATTAACAGGATCACTTTACGAGAGAAACAGGAGG 4728

Db 4321 CTGGAGTCTCTCTGCTGTCAAGTGAAGATATTAAGGATCACTTACGGAGAAACAGGAGG 4380  
Qy 4729 AAATAGCCCTCTCCAGGAGTTCACTGTGCTCTGGAGCAAGTCTACAGCTACCATCAGCGG 4788  
Db 4381 AAATAGCCCTCTCCAGGAGTTCACTGTGCTCTGGAGCAAGTCTACAGCTACCATCAGCGG 4440  
Qy 4789 CTTTAACTCGAGTTGATTAACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAG 4848  
Db 4441 CTTTAACTCGAGTTGATTAACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAG 4500  
Qy 4849 CCCGCAAGCAGCAGCCAAATTTCCATTAATTTACCGAACAGAAATTTGACAAACCATCCCA 4908  
Db 4501 CCCGCAAGCAGCAGCCAAATTTCCATTAATTTACCGAACAGAAATTTGACAAACCATCCCA 4560  
Qy 4909 GATGCAAGTGAAGGATGTTTCAAGCAACAGCAATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4968  
Db 4561 GATGCAAGTGAAGGATGTTTCAAGCAACAGCAATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4620  
Qy 4969 CCGTGTACTGTTACAGAGTAAGCCACCACTCCCAAAATTTGGACAGGACCAACAAAC 5028  
Db 4621 CCGTGTACTGTTACAGAGTAAGCCACCACTCCCAAAATTTGGACAGGACCAACAAAC 4680  
Qy 5029 TAAATCTCAGCTCAGATCAAAACAGAAATGACTATTGAAGGCTTGAGCCGACAGTGA 5088  
Db 4681 TAAATCTCAGCTCAGATCAAAACAGAAATGACTATTGAAGGCTTGAGCCGACAGTGA 4740  
Qy 5089 GTATGTGTTAGTGTCTATGCTCAGAAATCCCAAGCGGAGAGTCAAGCTCTGGTTCAAG 5148  
Db 4741 GTATGTGTTAGTGTCTATGCTCAGAAATCCCAAGCGGAGAGTCAAGCTCTGGTTCAAG 4800  
Qy 5149 TGCAGTAACCAATGATGATGCTTAAAGGACTGGCAATCACTGATGTGAGTFCGATTC 5208  
Db 4801 TGCAGTAACCAATGATGATGCTTAAAGGACTGGCAATCACTGATGTGAGTFCGATTC 4860  
Qy 5209 CATCAAAATTCCTGGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGTGACCTACTC 5268  
Db 4861 CATCAAAATTCCTGGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGTGACCTACTC 4920  
Qy 5269 GAGCCCTGAGGATGAAATCCATGAGCTATTCCTGCACTGATGATGATGATGATGATG 5328  
Db 4921 GAGCCCTGAGGATGAAATCCATGAGCTATTCCTGCACTGATGATGATGATGATGATG 4980  
Qy 5329 AGAGCTCAAGGCTCAGACCGGTTCTGAGTACACAGTCAAGTGTGCTGCTGACGA 5388  
Db 4981 AGAGCTCAAGGCTCAGACCGGTTCTGAGTACACAGTCAAGTGTGCTGCTGACGA 5040  
Qy 5389 TGATATGAGAGCCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCTGCAACCACTGA 5448  
Db 5041 TGATATGAGAGCCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCTGCAACCACTGA 5100  
Qy 5449 CTTGAAGTTCACTCAGTGTACACCCACAGGCTGAGGCCCCAGTGGACACACCCCAATGT 5508  
Db 5101 CTTGAAGTTCACTCAGTGTACACCCACAGGCTGAGGCCCCAGTGGACACACCCCAATGT 5160  
Qy 5509 TCAGCTCACTGATATCAGTGTGCGGTTGACCCCAAGGAGAGACCCGACCAATGAAGA 5568  
Db 5161 TCAGCTCACTGATATCAGTGTGCGGTTGACCCCAAGGAGAGACCCGACCAATGAAGA 5220  
Qy 5569 AATCAACCTTGCTCTGACAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCACCAA 5628  
Db 5221 AATCAACCTTGCTCTGACAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCACCAA 5280  
Qy 5629 ATATGAAGTGAAGTGTATGCTTTAAGGACACTTTGAACAGCAGACAGCTCAGGGAGT 5688  
Db 5281 ATATGAAGTGAAGTGTATGCTTTAAGGACACTTTGAACAGCAGACAGCTCAGGGAGT 5340  
Qy 5689 TGTCAACCTCTGGAGAAATGTCAGCCCAACAGAAAGGCTCTGTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCAACCTCTGGAGAAATGTCAGCCCAACAGAAAGGCTCTGTGTGACAGATGCTACTGA 5400  
Qy 5749 GACCAACCTCACTTGTGAGAAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5808  
Db 5401 GACCAACCTCACTTGTGAGAAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5460

Qy 5809 TCCCTTCCAGCCAAATGCGCAGATCTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5868  
Db 5461 TCCCTTCCAGCCAAATGCGCAGATCTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAG 5520  
Qy 5869 CTACACCATCACTGCTTACAAACAGGCACTGACTTCAAGATCTTACCTGTACACCTTGA 5928  
Db 5521 CTACACCATCACTGCTTACAAACAGGCACTGACTTCAAGATCTTACCTGTACACCTTGA 5580  
Qy 5929 TGACAAATGCTCGGAGCTCCCTGTGCTCATCGACGCTTCCACTGCTGCTGATGACCAATC 5988  
Db 5581 TGACAAATGCTCGGAGCTCCCTGTGCTCATCGACGCTTCCACTGCTGCTGATGACCAATC 5640  
Qy 5989 CAACCTGCTGCTTCCCTGCGCCACCAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 6048  
Db 5641 CAACCTGCTGCTTCCCTGCGCCACCAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 5700  
Qy 6049 TGCCAGGATTAACGCTACATCATCAAGTATGAGAAAGCTGCTGCTTCTCCAGAGAAATG 6108  
Db 5701 TGCCAGGATTAACGCTACATCATCAAGTATGAGAAAGCTGCTGCTTCTCCAGAGAAATG 5760  
Qy 6109 GGTCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6168  
Db 5761 GGTCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5820  
Qy 6169 CGAATATACAAATTTATGCTATTTGCTTCCAGAGTGAAGAAATTAATCAGAGAGCGAGCCCTGAT 6228  
Db 5821 CGAATATACAAATTTATGCTATTTGCTTCCAGAGTGAAGAAATTAATCAGAGAGCGAGCCCTGAT 5880  
Qy 6229 AAGGAAAGAGCAGAGGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6288  
Db 5881 AAGGAAAGAGCAGAGGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5940  
Qy 6289 ACCAGAGATCTTGGAGTGTCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6348  
Db 5941 ACCAGAGATCTTGGAGTGTCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6000  
Qy 6349 GTATGACCTGAAATGATTTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6408  
Db 6001 GTATGACCTGAAATGATTTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060  
Qy 6409 GCAACAAATGATCTTGGAGAACTGCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6468  
Db 6061 GCAACAAATGATCTTGGAGAACTGCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6120  
Qy 6469 CCCATAGGATGATGCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6528  
Db 6121 CCCATAGGATGATGCTTCCAGAGTTCACAAAGACCCCTTTCGTCAACCCCTGG 6180  
Qy 6529 GACAACTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6588  
Db 6181 GACAACTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6240  
Qy 6589 TGTGCTCACTGATGAGAAACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6648  
Db 6241 TGTGCTCACTGATGAGAAACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6300  
Qy 6649 TGTGCTCACTGATGAGAAACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6708  
Db 6301 TGTGCTCACTGATGAGAAACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360  
Qy 6709 GCAGAGGCTAAGGTTTGGGAGAGGTTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6768  
Db 6361 GCAGAGGCTAAGGTTTGGGAGAGGTTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6420  
Qy 6769 GACCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6828  
Db 6421 GACCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6480  
Qy 6829 AGATGAGTGGGAAACGAAATGCTGTAATCAGGCTTTTAACTGTTGCTGCTGCTGCTGCT 6888  
Db 6481 AGATGAGTGGGAAACGAAATGCTGTAATCAGGCTTTTAACTGTTGCTGCTGCTGCTGCT 6540











QY 5569 AATCAACCTTCTCTGACAGCTCATCGTGGTGTATCAGGACTTATGTTGGCCACCA 5628  
Db 5221 AATCAACCTTCTCTGACAGCTCATCGTGGTGTATCAGGACTTATGTTGGCCACCA 5280  
QY 5629 ATATGAAGTGAAGTGTATGCTCTTAAGGACACCTTTGACAGCAGACAGCTCAGGGAGT 5688  
Db 5281 ATATGAAGTGAAGTGTATGCTCTTAAGGACACCTTTGACAGCAGACAGCTCAGGGAGT 5340  
QY 5689 TGTCAACCTCTGAGAAATGTTCAGCCCAACGAGAGGGCTCGTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCAACCTCTGAGAAATGTTCAGCCCAACGAGAGGGCTCGTGTGACAGATGCTACTGA 5400  
QY 5749 GACCAACCTACCAATTAGCTGGAGAACCAAGACTGAGACCAATCACTGGCTTCCAAGTTGA 5808  
Db 5401 GACCAACCTACCAATTAGCTGGAGAACCAAGACTGAGACCAATCACTGGCTTCCAAGTTGA 5460  
QY 5809 TGCCTTCCAGCAATGCGCAGATCCCAATCCAGAGAACCAATCCAGCCAGATGTCAGAG 5868  
Db 5461 TGCCTTCCAGCAATGCGCAGATCCCAATCCAGAGAACCAATCCAGCCAGATGTCAGAG 5520  
QY 5869 CTACACCACTCACTGGCTTACCAACAGGCACTGACTTACAGATCTACCTGTACACCTTGA 5928  
Db 5521 CTACACCACTCACTGGCTTACCAACAGGCACTGACTTACAGATCTACCTGTACACCTTGA 5580  
QY 5929 TGCAATGCTCGGAGCTCCCTGTGTGTCATCGAGCCTTCCACTGCCATGATGACCAATC 5988  
Db 5581 TGAATGCTCGGAGCTCCCTGTGTGTCATCGAGCCTTCCACTGCCATGATGACCAATC 5640  
QY 5989 CAACCTGGTTCCTGGCCACACACCAATTCCTGTGCTGTATCATGCGACCGCCACG 6048  
Db 5641 CAACCTGGTTCCTGGCCACACACCAATTCCTGTGCTGTATCATGCGACCGCCACG 5700  
QY 6049 TGCCAGGATTAACCGCTACATCATCAAGTATGAGAAGCCTGGTCTCTCCAGAGAGT 6108  
Db 5701 TGCCAGGATTAACCGCTACATCATCAAGTATGAGAAGCCTGGTCTCTCCAGAGAGT 5760  
QY 6109 GGTCTCTGGCCCGCTGGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAAC 6168  
Db 5761 GGTCTCTGGCCCGCTGGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAAC 5820  
QY 6169 CGAATATCAATTTATGTCATTGCTCCCTGAAGAATAATCAGAGAGCGAGCCCTGATTGG 6228  
Db 5821 CGAATATCAATTTATGTCATTGCTCCCTGAAGAATAATCAGAGAGCGAGCCCTGATTGG 5880  
QY 6229 AAGGAAAAGACAGAGAGCTTCCCACTGGTAACTGTTCCACACCCCAATCTTCATGG 6288  
Db 5881 AAGGAAAAGACAGAGAGCTTCCCACTGGTAACTGTTCCACACCCCAATCTTCATGG 5940  
QY 6289 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCTGG 6348  
Db 5941 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCTGG 6000  
QY 6349 GTATGACACTGGAATGTTATCAGCTTCTGGCACTTCTGGTACGACACCCAGTGTGG 6408  
Db 6001 GTATGACACTGGAATGTTATCAGCTTCTGGCACTTCTGGTACGACACCCAGTGTGG 6060  
QY 6409 GCACAAATGATCTTTGAGGAACATGTTTAAAGCGGACCAACCGCCCAACAGGGCCAC 6468  
Db 6061 GCACAAATGATCTTTGAGGAACATGTTTAAAGCGGACCAACCGCCCAACAGGGCCAC 6120  
QY 6469 CCCCATAAGGCATAGGCCAAGACCAATACCGCCGAATGTAGGACAGAGCTCTCTCTCA 6528  
Db 6121 CCCCATAAGGCATAGGCCAAGACCAATACCGCCGAATGTAGGACAGAGCTCTCTCTCA 6180  
QY 6529 GACAACTCTCATGGGCCCAATTCAGGACACTTCTGAGTACATCATTTTCATGTCTCC 6588  
Db 6181 GACAACTCTCATGGGCCCAATTCAGGACACTTCTGAGTACATCATTTTCATGTCTCC 6240  
QY 6589 TGTGGCACTGATGAAGAACCTTTACGTTACAGGTTCTCGAATCTTACAGTGGCCAC 6648  
Db 6241 TGTGGCACTGATGAAGAACCTTTACGTTACAGGTTCTCGAATCTTACAGTGGCCAC 6300

QY 6649 TCTGACAGCCTCACAGAGTGGCCACCTTACAACTCATAGTGGAGGCACTGAAAGACCA 6708  
Db 6301 TCTGACAGCCTCACAGAGTGGCCACCTTACAACTCATAGTGGAGGCACTGAAAGACCA 6360  
QY 6709 GCAGAGGCATAAGGTTCCGGAGAGAGTGTGTACCGTGGCAACTCTGTCAACGAAGGCTT 6768  
Db 6361 GCAGAGGCATAAGGTTCCGGAGAGAGTGTGTACCGTGGCAACTCTGTCAACGAAGGCTT 6420  
QY 6769 GAACCAACTACGGATGACTCGTGTGTGACCCCTTACACAGTGTCCCAATTTATGCCGTTGG 6828  
Db 6421 GAACCAACTACGGATGACTCGTGTGTGACCCCTTACACAGTGTCCCAATTTATGCCGTTGG 6480  
QY 6829 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTTGTGCCAGTGTCTTAAAGCTT 6888  
Db 6481 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTTGTGCCAGTGTCTTAAAGCTT 6540  
QY 6889 TGSAAAGTGTCTATTTTCAGATGTGATTCATCTAGATGTGTGCCATGACAATGGTGTCAACTA 6948  
Db 6541 TGSAAAGTGTCTATTTTCAGATGTGATTCATCTAGATGTGTGCCATGACAATGGTGTCAACTA 6600  
QY 6949 CAAGATTTGAGAGAGTGGGACCGTCAAGGAGAAAATGGCCAGATGATGAGCTGCACATG 7008  
Db 6601 CAAGATTTGAGAGAGTGGGACCGTCAAGGAGAAAATGGCCAGATGATGAGCTGCACATG 6660  
QY 7009 TCTTGGGAACGAAAGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTTGTACGATGA 7068  
Db 6661 TCTTGGGAACGAAAGAGAAATCAAGTGTGACCCCTCATGAGGCAACGTTGTACGATGA 6720  
QY 7069 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 7128  
Db 6721 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 6780  
QY 7129 CTGCACTCTTTTGGAGCCAGCGGGCTGGCGCTGTGACACTGCCAGACCTGGGGG 7188  
Db 6781 CTGCACTCTTTTGGAGCCAGCGGGCTGGCGCTGTGACACTGCCAGACCTGGGGG 6840  
QY 7189 TGAACCCAGTCCCGAAGGCACTACTTGGCCAGTCTTACAAACCAAGTATTTCTCAGAGATACCA 7248  
Db 6841 TGAACCCAGTCCCGAAGGCACTACTTGGCCAGTCTTACAAACCAAGTATTTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAACAAACATTAATGTTAAATGGCCAAATGAGTGTCTTATGCCCTTTAGATGTACA 7308  
Db 6901 TCAGAGAACAAACATTAATGTTAAATGGCCAAATGAGTGTCTTATGCCCTTTAGATGTACA 6960  
QY 7309 GGTGTACAGAGAGATTTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAAGCATGT 7368  
Db 6961 GGTGTACAGAGAGATTTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAAGCATGT 7020  
QY 7369 CTCTGTGCCAAGATCCATCTAACTGGAGTGTAGTGTAGCAGACCCAGCTTAGAGTCTTC 7428  
Db 7021 CTCTGTGCCAAGATCCATCTAACTGGAGTGTAGTGTAGCAGACCCAGCTTAGAGTCTTC 7080  
QY 7429 TTTCTTTCTTAAGCCCTTTTGTCTGAGGAAATCTTCCAGCTTCCAGCTCAACTCAGCT 7488  
Db 7081 TTTCTTTCTTAAGCCCTTTTGTCTGAGGAAATCTTCCAGCTTCCAGCTCAACTCAGCT 7140  
QY 7489 TCTCCAAAGCATCACCTCGGAGTTCCTGAGGGTTTCTCAATAATGAGGGCTGCACATTT 7548  
Db 7141 TCTCCAAAGCATCACCTCGGAGTTCCTGAGGGTTTCTCAATAATGAGGGCTGCACATTT 7200  
QY 7549 GCTGTCTGTCTGAGGATTTCAATACCGCTCAGTATTTTAAATGAAGTGTCTTAAGA 7608  
Db 7201 GCTGTCTGTCTGAGGATTTCAATACCGCTCAGTATTTTAAATGAAGTGTCTTAAGA 7260  
QY 7609 TTTGGTTGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAATGTTTGAAT 7668  
Db 7261 TTTGGTTGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAATGTTTGAAT 7320  
QY 7669 GATATGACAAAATTTTAAGTAGGAAAGTCCACCAACACTTCTGCTTTCACTTAAGTGT 7728  
Db 7321 GATATGACAAAATTTTAAGTAGGAAAGTCCACCAACACTTCTGCTTTCACTTAAGTGT 7380  
QY 7729 CTGCCCCGCAATCTGTAGGAAACAAGCATGATCTTGTACTGTGATATTTTAAATATCA 7788

```

181 CGAAGTAAACCTGAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCG 240
589 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTCTCATCGG 648
241 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTCTCATCGG 300
649 GGTGGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCATGCAAGGGGTGAGTC 708
301 GGTGGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCATGCAAGGGGTGAGTC 360
709 CTACAAGATTGCTGACACCTGAGGAGACCACTGAGACTGGTGGTTACATGTTAGAGTG 768
361 CTACAAGATTGCTGACACCTGAGGAGACCACTGAGACTGGTGGTTACATGTTAGAGTG 420
769 TGTGTGCTTCTGTAATGGAAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAAAGTGT 828
421 TGTGTGCTTCTGTAATGGAAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAAAGTGT 480
829 TGATCATGCTGCTGGGACTCCCTATGTTGGTGGGAGAAACCTGGGAGAACCTTACCAGG 888
481 TGATCATGCTGCTGGGACTCCCTATGTTGGTGGGAGAAACCTGGGAGAACCTTACCAGG 540
889 CTGGATGATGTAGATTGTACTTTCCTGGGAGAAAGGAGCGGACGATCACTTGCATTC 948
541 CTGGATGATGTAGATTGTACTTTCCTGGGAGAAAGGAGCGGACGATCACTTGCATTC 600
949 TAGAANTAGATGCAACGATCAGGACACAGGACATCTATAGATTGGAGACACTGGAG 1008
601 TAGAANTAGATGCAACGATCAGGACACAGGACATCTATAGATTGGAGACACTGGAG 660
1009 CAAGAAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCCGAGGAGA 1068
661 CAAGAAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCCGAGGAGA 720
1069 GTGAAGTGTGAGAGGACACACTCTGTGAGAGACACATCGAGCGGATCTGGCCCTTCAC 1128
721 GTGAAGTGTGAGAGGACACACTCTGTGAGAGACACATCGAGCGGATCTGGCCCTTCAC 780
1129 CGATGTTGCTGAGCTGTTTACCAACCGGACGCTCACCCTCCAGGCTCTCTCCCTATGGCCA 1188
781 CGATGTTGCTGAGCTGTTTACCAACCGGACGCTCACCCTCCAGGCTCTCTCCCTATGGCCA 840
1189 CTGTGTACAGACAGTGGTGTGTTACTCTGTGGGAGTGCAGTGGCTGAGACACAGG 1248
841 CTGTGTACAGACAGTGGTGTGTTACTCTGTGGGAGTGCAGTGGTGGAGACACAGG 900
1249 AAATAAGCAATGCTTGTGACGTCCTGGGCAACCGGAGTGCAGTGGTGGAGACAGCTGT 1308
901 AAATAAGCAATGCTTGTGACGTCCTGGGCAACCGGAGTGCAGTGGTGGAGACAGCTGT 960
1309 AACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTGTCTTACATTCACCTACAA 1368
961 AACCCAGACTTACGGTGGCAACTTAAATGGAGAGCCATGTGTCTTACATTCACCTACAA 1020
1369 TGGCAGGAGCTTCTACTCTGCACACAGAGGCGGACAGGAGCATCTTTGGTGCAG 1428
1021 TGGCAGGAGCTTCTACTCTGCACACAGGAGGCGGACAGGAGCATCTTTGGTGCAG 1080
1429 CACAACCTTGAATTTATGAGCAGGACCAAGAAATCTCTTTCTGCACAGACCACTGT 1488
1081 CACAACCTTGAATTTATGAGCAGGACCAAGAAATCTCTTTCTGCACAGACCACTGT 1140
1489 GTTTCAGACTCGAGAGGAAATTCGAATGGTGGCTGTGGCACTTCCCTCCCTATACAA 1548
1141 GTTTCAGACTCGAGAGGAAATTCGAATGGTGGCTGTGGCACTTCCCTCCCTATACAA 1200
1549 CAACCAATTTACACTGATTGCACTTCTGAGGGGAGAGAGACAAATGAAGTGGTGG 1608
1201 CAACCAATTTACACTGATTGCACTTCTGAGGGGAGAGAGACAAATGAAGTGGTGG 1260
1609 GACCACAGAACTATGATGCCGACAGAGTTTGGGTTCTGCCCCCATGGTCCCCACGA 1668
1261 GACCACAGAACTATGATGCCGACAGAGTTTGGGTTCTGCCCCCATGGTCCCCACGA 1320

```

7381 CTGCGCGCAATACTGTAGGAACAAGCATGATCTGTACTGTGATATTTAAATATCCA 7440

7789 CAGTACT 7795

7441 CAGTACT 7447

RESULT 15

FOI1162

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PR

PI

PC

C07K15/12,

PC

C12N15/00,

C12P21/00,

C12R1:865;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

[illegible]

...





QY	6049	TGCCAGGATTACCGCTACATCATCAAGTATGAGAAGCTGGGCTCTCTCCACAGAGAAT	6108
Db	5701	TGCCAGGATTACCGCTACATCATCAAGTATGAGAAGCTGGGCTCTCTCCACAGAGAAT	5760
QY	6109	GGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTATTACTGGCTGTGAACCGGGAAC	6168
Db	5761	GGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTATTACTGGCTGTGAACCGGGAAC	5820
QY	6169	CGAATATACAAATTTATGTCAATTCCTCAAGAAATAATCAGAAAGAGAGCCCTTGATTGG	6228
Db	5821	CGAATATACAAATTTATGTCAATTCCTCAAGAAATAATCAGAAAGAGAGCCCTTGATTGG	5980
QY	6229	AAGGAAAAGACAGACGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGG	6288
Db	5881	AAGGAAAAGACAGACGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGG	5940
QY	6289	ACCAGAGATCTTGGATGTTCTTCCACAGTTTCAAAAGACCCCTTTCGTACCCACCCCTGG	6348
Db	5941	ACCAGAGATCTTGGATGTTCTTCCACAGTTTCAAAAGACCCCTTTCGTACCCACCCCTGG	6000
QY	6349	GTATGACACTGGAAATGGTATTACGCTTCTGGCACTTCTGGTCAGCAACCCAGTGTGG	6408
Db	6001	GTATGACACTGGAAATGGTATTACGCTTCTGGCACTTCTGGTCAGCAACCCAGTGTGG	6060
QY	6409	GCAACAAATGATCTTTGAGGAACATGGTTTAGCGGACCAACCGCCACAAAGCGCCAC	6468
Db	6061	GCAACAAATGATCTTTGAGGAACATGGTTTAGCGGACCAACCGCCACAAAGCGCCAC	6120
QY	6469	CCCCATAAGGCATAGGCCAAAGACCAATACCGCCGAATGTAGGCAAGAAAGCTCTCTCA	6528
Db	6121	CCCCATAAGGCATAGGCCAAAGACCAATACCGCCGAATGTAGGCAAGAAAGCTCTCTCA	6180
QY	6529	GACAAACATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCATTTTCTATCTATCC	6588
Db	6181	GACAAACATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCATTTTCTATCTATCC	6240
QY	6589	TGTTGGCACTGATGAAGAACCCCTTACAGTTTCTGGAACTTCTACAGTGGCCAC	6648
Db	6241	TGTTGGCACTGATGAAGAACCCCTTACAGTTTCTGGAACTTCTACAGTGGCCAC	6300
QY	6649	TCGACAGCCCTCAACAGAGGTGCCACCTCAACATCATATGAGGAGGCACTGAAGACCA	6708
Db	6301	TCGACAGCCCTCAACAGAGGTGCCACCTCAACATCATATGAGGAGGCACTGAAGACCA	6360
QY	6709	GCAGAGGCATAGGTTTCGGGAAGAGGTTCTTACCGTGGCACTCTGTCAACGAGGCTT	6768
Db	6361	GCAGAGGCATAGGTTTCGGGAAGAGGTTCTTACCGTGGCACTCTGTCAACGAGGCTT	6420
QY	6769	GAAACCACTACCGATGACTCGTGTGTTGACCCCTACACAGTGTCCCATTTATGCGGTTGG	6828
Db	6421	GAAACCACTACCGATGACTCGTGTGTTGACCCCTACACAGTGTCCCATTTATGCGGTTGG	6480
QY	6829	AGATGAGTGGGACGAATCTCTGAATCAGGCTTTAAACCTGTTGCCAGTGTAGGCTT	6888
Db	6481	AGATGAGTGGGACGAATCTCTGAATCAGGCTTTAAACCTGTTGCCAGTGTAGGCTT	6540
QY	6889	TGGAAGTGTCTATTTCAGATGTGATTCTATAGATGGTGGCCATGACATTTGGTGAACCTA	6948
Db	6541	TGGAAGTGTCTATTTCAGATGTGATTCTATAGATGGTGGCCATGACATTTGGTGAACCTA	6600
QY	6949	CAAGATTGAGAGAGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACATG	7008
Db	6601	CAAGATTGAGAGAGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACATG	6660
QY	7009	TCCTTGGGAACGGAAAAGGAGAAATTCAGTGTGACCCCTCATATGAGGCAACGTTTACGATGA	7068
Db	6661	TCCTTGGGAACGGAAAAGGAGAAATTCAGTGTGACCCCTCATATGAGGCAACGTTTACGATGA	6720
QY	7069	TGGGAAGACATACAGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC	7128
Db	6721	TGGGAAGACATACAGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC	6780

QY	7129	CTGCACATGCTTTTGGAGCCGACGCGGGCTGGCGCTGTGACAACTGCCGACACCTTGGGG	7188
Db	6781	CTGCACATGCTTTTGGAGCCGACGCGGGCTGGCGCTGTGACAACTGCCGACACCTTGGGG	6840
QY	7189	TGAACCCGATCCCGAAGGACATCTACTGCGCAGTCTTACAAACAGATATTCTCAGAGATACA	7248
Db	6841	TGAACCCGATCCCGAAGGACATCTACTGCGCAGTCTTACAAACAGATATTCTCAGAGATACA	6900
QY	7249	TCAGAGAAACAAACACTAATGTTTAAATGGCCCAATTTAGTGTCTTCAATGCTTTAGATGTA	7308
Db	6901	TCAGAGAAACAAACACTAATGTTTAAATGGCCCAATTTAGTGTCTTCAATGCTTTAGATGTA	6960
QY	7309	GGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTCCAAATCCAGAGAGAAACAAACATGT	7368
Db	6961	GGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTCCAAATCCAGAGAGAAACAAACATGT	7020
QY	7369	CTCTCTGCCAAGATCCATCTAAACTGGAGTGTATTTAGCAGACCCAGCTTAGAGTCTTTC	7428
Db	7021	CTCTCTGCCAAGATCCATCTAAACTGGAGTGTATTTAGCAGACCCAGCTTAGAGTCTTTC	7080
QY	7429	TTTCTTTCTTAAAGCCCTTTGCTCTGGAGGAGTCTTCCAGCTTCCAGCTTCAACTCAGCT	7488
Db	7081	TTTCTTTCTTAAAGCCCTTTGCTCTGGAGGAGTCTTCCAGCTTCCAGCTTCAACTCAGCT	7140
QY	7489	TTCTCAAGCATCACCTCTGGAGTCTTCTGAGGGTTTCTCATAAATGAGGGCTGCACATT	7548
Db	7141	TTCTCAAGCATCACCTCTGGAGTCTTCTGAGGGTTTCTCATAAATGAGGGCTGCACATT	7200
QY	7549	GCCTGTTCTGCTTGGAGTATTCAATACCGCTCAGTATTTTAAATGAAAGTGTCTTAAAG	7608
Db	7201	GCCTGTTCTGCTTGGAGTATTCAATACCGCTCAGTATTTTAAATGAAAGTGTCTTAAAG	7260
QY	7609	TTTGGTTTGGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAATGTTTGAAT	7668
Db	7261	TTTGGTTTGGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAATGTTTGAAT	7320
QY	7669	GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCCAAAGACTTCTGCTTTCACCTTAAAGT	7728
Db	7321	GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCCAAAGACTTCTGCTTTCACCTTAAAGT	7380
QY	7729	CTGCGCCGCAATCTCTAGGAAACCAAGCATGATCTTGTACTGTGATATTTTAAATATCCA	7788
Db	7381	CTGCGCCGCAATCTCTAGGAAACCAAGCATGATCTTGTACTGTGATATTTTAAATATCCA	7440
QY	7789	CAGTACT 7795	
Db	7441	CAGTACT 7447	

Search completed: March 7, 2004, 16:57:10  
Job time : 28686 secs









QY 536 CTGGGAGAGAGAAACCCAGGCTGACCCAGGCTCTTTTTCAGGAGATGTGGAGG 595  
 Db 543 CTGGGAGAGAGAAACCCAGGCTGACCCAGGCTCTTTTTCAGGAGATGTGGAGG 602  
 QY 556 CATCTCCAGGCCATCTGGCTGGTGGTGAACACACGAGAGTGTCTGACATCCAGTGGT 655  
 Db 603 CATCTCCAGGCCATCTGGCTGGTGGTGAACACACGAGAGTGTCTGACATCCAGTGGT 662  
 QY 656 TCGGAGAGCCCTATGGGCGGTGACGACAGCGTCCGGTGTGACGTTGAGAGAGCCG 715  
 Db 663 TCGGAGAGCCCTATGGGCGGTGACGACAGCGTCCGGTGTGACGTTGAGAGAGCCG 722  
 QY 716 ACAGCAGCGGCGTGGTGGTTCACGCCAGCGGTGACGATCTGAGTCAGAAATGTGGCCT 775  
 Db 723 ACAGCAGCGGCGTGGTGGTTCACGCCAGCGGTGACGATCTGAGTCAGAAATGTGGCCT 782  
 QY 776 GGACAACTTCGGGAGCTTTGACTGGGTGATCATGAGAAACATGAGTTGAAACAGCGCTGGA 835  
 Db 783 GGACAACTTCGGGAGCTTTGACTGGGTGATCATGAGAAACATGAGTTGAAACAGCGCTGGA 842  
 QY 836 GGAGCAGTTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAG 895  
 Db 843 GGAGCAGTTGGAG-AVCTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAG 901  
 QY 896 GAGTGAGCTGGGCGCTGCTGAGTGGGGTGGGCTGACTCTGCAAAATGGGGTGTCC 955  
 Db 902 GAGTGAGCTGGGCGCTGCTGAGTGGGGTGGGCTGACTCTGCAAAATGGGGTGT-CC 960  
 QY 956 CCGATCTCTGGCGGAGGTGAGGAAACAGACAGCGGGGCTGAGATCTCAGGGGTTGT 1013  
 Db 961 CCGATCTCTGGCGGAGGTGAGGAAACAGGAGAGGGGCTGATATGTGAGGGTGTGAT 1018

RESULT 4  
 3M808695  
 LOCUS  
 DEFINITION AGNCOURT 6582590 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5471256  
 5', mRNA sequence.

ACCESSION BM808695  
 VERSION  
 KEYWORDS  
 SOURCE EST.

ORGANISM Homo sapiens (human)  
 Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1080)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCMI979 row: 9 column: 01  
 High quality sequence start: 15  
 High quality sequence stop: 695.  
 Location/Qualifiers

FEATURES  
 source

1. .1080  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5471256"  
 /issue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_41"  
 /notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:

# ORIGIN

Query Match 69.0%; Score 884.4; DB 12; Length 1080;  
 Best Local Similarity 97.4%; Pred. No. 6.5e-200;  
 Matches 962; Conservative 0; Mismatches 18; Indels 8; Gaps 6;

QY 302 CATGCCCGCGCTGGGAGCGCGCCCGCGCTGTACTGTCTTTCAGCGCAAGAGAAATC 361  
 Db 27 CATGCCCGCGCTGGGATGGCGCCCGCGCTGTACTGTCTTTCAGCGCAAGAGAAATC 86  
 QY 362 CGGGAAGAGATTCGTGACGAGAGCGCTGACAGAGAGATTCGAGCTGATGTCTGTGT 421  
 Db 87 CGGGAAGAGATTCGTGACGAGAGCGCTGACAGAGAGATTCGAGCTGATGTCTGTGT 146  
 QY 422 CTTCCGGCTCTCTGTCTCACTCAAGGAACAGTATGCTCAGGAGCATGCTTGAACCTCCA 481  
 Db 147 CTTCCGGCTCTCTGTCTCACTCAAGGAACAGTATGCTCAGGAGCATGCTTGAACCTCCA 206  
 QY 482 GAGACTCTCTGACACACGACACTCAAGGAGCGCTTTCGGAAGGACATGATCCCGTGGG 541  
 Db 207 GAGACTCTCTGACACACGACACTCAAGGAGCGCTTTCGGAAGGACATGATCCCGTGGG 266  
 QY 542 AGAGAGAGAAACGCCAGGCTGACCCAGGCTTCTTTGAGGAGATTTGAGGAGCATCTC 601  
 Db 267 AGAGAGAGAAACGCCAGGCTGACCCAGGCTTCTTTGAGGAGATTTGAGGAGCATCTC 326  
 QY 602 CCAGCCCATCTGGCTGTGAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGGA 661  
 Db 327 CCAGCCCATCTGGCTGTGAGTGACACACGAGAGTGTCTGACATCCAGTGGTTTCGGGA 386  
 QY 662 GGCCTATCGGGCGCTGACGACAGAGCGTCCGGTGTGAGCGTTGAGGAGCGCCACAGCA 721  
 Db 387 GGCCTATCGGGCGCTGACGACAGAGCGTCCGGTGTGAGCGTTGAGGAGCGCCACAGCA 446  
 QY 722 GCGGGCTGGGTGTTTACGCCAGGGGTGGACGATGCTGAGTCAGAAATGTGGCTTGGCAA 781  
 Db 447 GCGGGCTGGGTGTTTACGCCAGGGGTGGACGATGCTGAGTCAGAAATGTGGCTTGGCAA 506  
 QY 782 CTTGGGAGCTTTCAGCTGGGTGATCGAGAACCTGAGTTGAACAGCGCCCTGGAGGAGCA 841  
 Db 507 CTTGGGAGCTTTCAGCTGGGTGATCGAGAACCTGAGTTGAACAGCGCCCTGGAGGAGCA 566  
 QY 842 GTTGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGA 901  
 Db 567 GTTGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGA 626  
 QY 902 GCTGGGCTCTCTGAGTGGGGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCCGATC 961  
 Db 627 GCTGGGCTCTCTGAGTGGGGTGGGCTGACTCTGCAAAATGGGGGTGTCCCCCGATC 686  
 QY 962 CTGGCCAGGTGAGGAAACAGACAGCGGGGTCTAGATTCTGAGGGGTGTGGTATATTG 1021  
 Db 687 CTGGCCAGGTGAGGAAACAGACAGCGGGGTCTAGATTCTGAGGGGTGTGGTATATTG 746  
 QY 1022 GGCAAGGCA-GGAAACCTCTGGAGACCTCATTTTCTCCATGGGGAACAGACCATGTCT 1080  
 Db 747 GGCAAGGCAAGGAAACCTCTGGAGACCTCATTTTCTCCATGGGGAACAGACCATGTCT 806  
 QY 1081 TCAGGAGGAGACTCCAAAGGCAAGGAGGGTGTCTTGGCTGTGCTTGAAGCGCAACCCCT 1140  
 Db 807 TCAGGAGGAGACTCNCANGGCAAGGAGGGTGTCTTGGCTGTGCTTGAAGCGCAACCCCT 866  
 QY 1141 GCCATATCCCGAGTGGCAGTCCCTCAGCCT-GTGGTGGCTTGGCA-TCTGACTGGATG 1198  
 Db 867 GCCATATCCCGAGTGGCAGTCCCTCAGCCTGGTGGTGGCTTGGCA-TCTGACTGGATG 926  
 QY 1199 TTTTCAGCCCC--TTGTTCTGGGCAAGAACCCAGAGCTCCCCAGT-GTGGATACTATAA 1255

GCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

```

2b 927 TTCTTCCGCCCCCTTGTCTCTGGGCAAGACCCAGAACCTCCCAAGTCTCCAGTGGTGGATACTAATA 986
1256 A--CCTCTTGAGACACAAAAA 1281
2b 987 AACCTCTTGAGCCCTCAAAAAA 1014

RESULT 5
LOCUS BUL67861
DEFINITION AGENCOURT_8050094 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6089091
5', mRNA sequence.
ACCESSION BUL67861
VERSION BUL67861.1 GI:22681845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2328 row: f column: 04
High quality sequence stop: 667.
Location/Qualifiers
1. .885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6089091"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 112"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

DRIGIN
Query Match 67.0%; Score 858.2; DB 13; Length 885;
Best Local Similarity 98.5%; Pred. No. 1.1e-193;
Matches 866; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

2y 51 ACAGAGGAGGAGCTCTCTGTGAAGAACCGCTTCTCTCCGGGTGTCTCACCTTTTCT 110
2b 1 ACAGAGGAGGAGCTCTCTGTGAAGAACCGCTTCTCTCCGGGTGTCTCACCTTTTCT 60

2y 111 CCCCATATCTGTTTGGACATGATGCTGAGGACAGGTCCGCGGGCGGTGAGCCCTGTTCCGA 170
2b 61 CCCCATATCTGTTTGGACATGATGCTGAGGACAGGTCCGCGGGCGGTGAGCCCTGTTCCGA 120

2y 171 GCTACGCGAGAGGCGGAGTGTCTTGTGCGGTCCGCTTAGTGGCGCGGTCCCA 230
2b 121 GCTACGCGAGAGGCGGAGTGTCTTGTGCGGTCCGCTTAGTGGCGCGGTCCCA 180

2y 231 TTCGCGCGGTGTCCTCCGATTTAGGGTAGGAGAGTGTACGCTTCAGGATCGGAGG 290
2b 181 TTCGCGCGGTGTCCTCCGATTTAGGGTAGGAGAGTGTACGCTTCAGGATCGGAGG 240

```

```

291 CTTGGCGCCCCCTGCGCCCTCTGGAGGCGCCCGGGGCTGGTACTGCTCTTTCAGCGGC 350
241 CTTGGCGCCCCCTGCGCCCTCTGGAGGCGCCCGGGGCTGGTACTGCTCTTTCAGCGGC 300
351 AAGAGGAAATCCGGGAAGGACTTGTGACCGAGCGCTGACAGCAGACTTTGGAGCTGAT 410
301 AAGAGGAAATCCGGGAAGGACTTGTGACCGAGCGCTGACAGCAGACTTTGGAGCTGAT 360
411 GTCTGTGTCTCTCCGCTCTCTGGTCCACTCAAGAAACAGTAGTATGCTCAGGAGCATGGC 470
361 GTCTGTGTCTCTCCGCTCTCTGGTCCACTCAAGAAACAGTAGTATGCTCAGGAGCATGGC 420
471 TTGAACCTTCCAGAGACTCTCTGGACACACAGCACTTACAGAGAGGCTTTTCGGAAGGACATG 530
421 TTGAACCTTCCAGAGACTCTCTGGACACACAGCACTTACAGAGAGGCTTTTCGGAAGGACATG 480
531 ATCCGCTGGGAGAGGAGAAACCGCAGCTGACCCAGGCTTCTTTTGAGGAAGATTGTG 590
481 ATCCGCTGGGAGAGGAGAAACCGCAGCTGACCCAGGCTTCTTTTGAGGAAGATTGTG 540
591 GAGGCACTCTCCAGGCCATCTGCTGTGTAGTGACACACAGAGAGTGTCTGACATCCAG 650
541 GAGGCACTCTCCAGGCCATCTGCTGTGTAGTGACACACAGAGAGTGTCTGACATCCAG 600
651 TGGTTTCGGGAGGCTATGGGCGCTGACGACAGCGTCCGCTTGTAGCGTTGGAGCAG 710
601 TGGTTTCGGGAGGCTATGGGCGCTGACGACAGCGTCCGCTTGTAGCGTTGGAGCAG 660
711 AGCCGACAGCAGCGGGGCTGGGTGTTTTCAGCCAGGGGTGACAGTGTCTGACATGATGT 770
661 AGCCGACAGCAGCGGGGCTGGGTGTTTTCAGCCAGGGGTGACAGTGTCTGACATGATGT 720
771 GCGCTGGACAACTTCGGGAGCTTTGACTGGGTGCTACGAGACCATGGAGTTGAACAGCGC 830
721 GCGCTGGACAACTTCGGGAGCTTTGACTGGGTGCTACGAGACCATGGAGTTGAACAGCGC 780
831 CTGGAGGAGCAGTGTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCTACATAGT 890
781 CTGGAGGAGCAGTGTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCTACATAGT 840
891 TCTAGAGTGTAGCTGGGCGCTGCTGAGGTGGGTGCTGAGTGGGTGAGTGAACAGCGC 929
841 TCTAGAGTGTAGCTGGGCGCTGCTGAGGTGGGTGCTGAGTGGGTGAGTGAACAGCGC 879

RESULT 6
BQ888524 309 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8071320 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6090682
5', mRNA sequence.
ACCESSION BQ888524
VERSION BQ888524.1 GI:22280538
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2332 row: h column: 11
High quality sequence stop: 641.
Location/Qualifiers

```

FEATURES

```

source
1. 909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6090682"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 112"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 65.2%; Score 834.8; DB 13; Length 909;
Best Local Similarity 97.7%; Pred. No. 4.2e-188;
Matches 868; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

QY 51 ACAGGGGAGGCTGCTCTGAGGAGGACCGCTTCTCTCGCGGTCTCACCCCTTTCT 110
DB 1 ACAGGGGAGGCTGCTCTGAGGAGGACCGCTTCTCTCGCGGTCTCACCCCTTTCT 60

QY 111 CCCATATCTGTTGGACATGAGTGAAGGACGCTCGCGCGGTCTACGCCCTGTTCGA 170
DB 61 CCCATATCTGTTGGACATGAGTGAAGGACGCTCGCGCGGTCTACGCCCTGTTCGA 120

QY 171 GCTAGGCGGAGGAGGCGCGGATGTTCTCTGCTCGCGGTCTGAGCGCGGTCCA 230
DB 121 GCTAGGCGGAGGAGGCGCGGATGTTCTCTGCTCGCGGTCTGAGCGCGGTCCA 180

QY 231 TTCGCGCGCGTGTCCGATTTTAGGGTAGGAGAGTGTACGCTTCAGGCATCCGAGG 290
DB 181 TTCGCGCGCGTGTCCGATTTTAGGGTAGGAGAGTGTACGCTTCAGGCATCCGAGG 240

QY 291 CGTGGCGGCGCCATGCGCCCGCTGGAGGCGCGCGCGCTGTACTGCTGTACGCGGC 350
DB 241 CGTGGCGGCGCCATGCGCCCGCTGGAGGCGCGCGCGCTGTACTGCTGTACGCGGC 300

QY 351 AAGAGGAAATCCGGGAGGAGACTTCGTGACCGAGGCGCTCAGAGCAGACTTTGGAGCTGAT 410
DB 301 AAGAGGAAATCCGGGAGGAGACTTCGTGACCGAGGCGCTCAGAGCAGACTTTGGAGCTGAT 360

QY 411 GTCTGTGCTGCTCCGGCTCTCTGTCTCACTCAAGGAACAGATGCTCAGAGCATGGC 470
DB 361 GTCTGTGCTGCTCCGGCTCTCTGTCTCACTCAAGGAACAGATGCTCAGAGCATGGC 420

QY 471 TTGAATCTCCAGAGACTCTCGGCTCTGTGCTCACTCAAGGAGGCGCTTCGGAAGGACATG 530
DB 421 TTGAATCTCCAGAGACTCTCGGCTCTGTGCTCACTCAAGGAGGCGCTTCGGAAGGACATG 480

QY 531 ATCCGCTGGGAGGAGAGAAACCCAGGCTGACCCAGGCTTTTTCAGGAGAGATTGTG 590
DB 481 ATCCGCTGGGAGGAGAGAAACCCAGGCTGACCCAGGCTTTTTCAGGAGAGATTGTG 540

QY 591 GAGGCGATCTCCAGGCGCATCTGCTGTGTGAGTACACAGGAGAGTGTCTGACATCCAG 650
DB 541 GAGGCGATCTCCAGGCGCATCTGCTGTGTGAGTACACAGGAGAGTGTCTGACATCCAG 600

QY 651 TGTTTTCGGAGGCGCTATGCGGCGGTGACGAGACGCTCCGCTTTGACGTTTGAGAGAG 710
DB 601 TGTTTTCGGAGGCGCTATGCGGCGGTGACGAGACGCTCCGCTTTGACGTTTGAGAGAG 660

QY 711 AGCCGACAGCAGCGGCGGTGTTTCAGCCAGGCGGTGAGATGCTCAGTCAGATGT 770
DB 661 AGCCGACAGCAGCGGCGGTGTTTCAGCCAGGCGGTGAGATGCTCAGTCAGATGT 720

QY 771 GGCTTGGACAACTTCGGGAGCTTTTGA CTGGTCTCATCGAGAACATGGAATTGAAACAGGCG 830
DB 721 GGCTTGGACAACTTCGGGAGCTTTTGA CTGGTCTCATCGAGAACATGGAATTGAAACAGGCG 780

QY 831 CTGGAGGAGCAGTTGGAGAACCTGTAGATTTATCCG-CTCAGACTTTAGTCACTA-- 887
DB 781 CTGGAGGAGCAGTTGGAGAACCTGTAGATTTATCCGCTCCAGACTTTAGTCACTAGG 840

QY 888 GGTCTTAGGAGTGAAGTGGGCGCTGCTGAGGTGGGGTGGGCGTGAAT 935
DB 841 GGTCTTAGGAGTGAAGTGGGCGCTGCTTTAAAGGGGGGGTGGGCGCT 888

RESULT 7
BQ057770
LOCUS
DEFINITION
BQ057770
ACCESSION
VERSION
BQ057770.1 GI:19817110
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 966)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-k@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2065 row: j column: 08
High quality sequence stop: 705.

FEATURES
source
1. 966
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5813479"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 65.2%; Score 834.6; DB 13; Length 966;
Best Local Similarity 98.2%; Pred. No. 4.8e-188;
Matches 865; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 403 GAGCTGATGCTGTGCTCTCTCGGCTCTCTCGCTCTCTCGCTCTCTCGCTCTCTCGCT 462
DB 1 GAGCTGATGCTGTGCTCTCTCGGCTCTCTCGCTCTCTCGCTCTCTCGCTCTCTCGCT 60

QY 463 AGCATGGCTTTGAATCTCCAGAGACTCTGGACACACGACCTTCAAGAGGCGCTTTCCGA 522
DB 61 AGCATGGCTTTGAATCTCCAGAGACTCTGGACACACGACCTTCAAGAGGCGCTTTCCGA 120

QY 523 AGGACATGATCCCTGGGAGAGAGAAACGCCAGGCTGACCCAGGCTCTTTTTCAGGA 582
DB 121 AGGACATGATCCCTGGGAGAGAGAAACGCCAGGCTGACCCAGGCTCTTTTTCAGGA 180

```



```

583 AGATTGTGGAGGCGATCTCCAGCCCATCTGGCTGGTGTGAGTGACACACGAGAGTGTCTG 642
Db 181 AGATTGTGGAGGCGATCTCCAGCCCATCTGGCTGGTGTGAGTGACACACGAGAGTGTCTG 240
2y 643 ACATCCAGTGGTTTCGAGAGCCCTATCGGCGCGTGTACGACAGACGGTCCGCTGTAGCGT 702
Db 241 ACATCCAGTGGTTTCGAGAGCCCTATCGGCGCGTGTACGACAGACGGTCCGCTGTAGCGT 300
2y 703 TGGAGCAGAGCCGACAGACAGCGGGCTGGGTGTTCACGCCAGAGGGTGGAGTGTGAGT 762
Db 301 TGGAGCAGAGCCGACAGACAGCGGGCTGGGTGTTCACGCCAGAGGGTGGAGTGTGAGT 360
2y 763 CAGAAATGTGGCTGGACAACTTCGGGACCTTGTGCTGGGTCACTCGAAGAACCATGGAGTTG 822
Db 361 CAGAAATGTGGCTGGACAACTTCGGGACCTTGTGCTGGGTCACTCGAAGAACCATGGAGTTG 420
2y 823 AACAGCGCTCGGAGAGCAGTTCGAGAACCTGTAGATTTATCCGCTCCAGACTTTAGT 882
Db 421 AACAGCGCTCGGAGAGCAGTTCGAGAACCTGTAGATTTATCCGCTCCAGACTTTAGT 480
2y 883 CACTAGGTTCTAGAGTGTGAGCTGGGGCTGTGTGAGGTGGGGTGGGCTGACTCTGCAAA 942
Db 481 CACTAGGTTCTAGAGTGTGAGCTGGGGCTGTGTGAGGTGGGGTGGGCTGACTCTGCAAA 540
2y 943 ATGGGGTGTCCCGGATCTTCGCGGAGGTGTGAGAACAGACAGAGGGGGTCTAGATCTG 1002
Db 541 ATGGGGTGTCCCGGATCTTCGCGGAGGTGTGAGAACAGACAGAGGGGGTCTAGATCTG 600
2y 1003 AGGGGTTGTGGATATTGGGCAAGGAGAGAACTCTGGAGACCTCATTTTCTCCATGG 1062
Db 601 AGGGGTTGTGGATATTGGGCAAGGAGAGAACTCTGGAGACCTCATTTTCTCCATGG 660
2y 1063 GGAAGACAGCCTGCTCTTCAGAGAGAGACTCCAGGGCAAGAGAGGGTGTCTGGCTGT 1122
Db 661 GGAAGACAGCCTGCTCTTCAGAGAGAGACTCCAGGGCAAGAGAGGGTGTCTGGCTGT 720
2y 1123 GCTTGAAGGCAAACTCTCCAGTATCCAGTGTCCAGTGTCCAGTGTCCAGTGTCCAGT 1182
Db 721 GCTTGAAGGCAAACTCTCCAGTATCCAGTGTCCAGTGTCCAGTGTCCAGTGTCCAGT 780
2y 1183 GCATCTGACTGTGATTTTCAGCCCTTGTTCAGCCCTTGTTCAGCCCTTGTTCAGCCCT 1241
Db 781 GCATCTGACTGTGATTTTCAGCCCTTGTTCAGCCCTTGTTCAGCCCTTGTTCAGCCCT 840
2y 1242 GT-CGATCTAATAAATCTCTGGAGCAGCAAAAAA 1281
Db 841 GTGGGAACTAAATAAATCTCTGGAGAGCCCAAAAAA 881

RESULT 8
LOCUS BUI47264
DEFINITION AGENCOURT_8070801 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6089738
5', mRNA sequence.
ACCESSION BUI47264
VERSION BUI47264.1 GI:22660796
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 888)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

```

http://image.llnl.gov
Plate: LLOW2330 row: a column: 03
High quality sequence stop: 698.
Location/Qualifiers
1..888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="IMAGE:6089738"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 65.0%; Score 833.2; DB 13; Length 888;
Best Local Similarity 98.2%; Pred. No. 1e-187;
Matches 864; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 51 ACAGGGGAGGCTGTCTGTGAAGAACCGCTTTCTTCGCGGTGTCTACCCCTTTCT 110
Db 1 ACAGGGGAGGCTGTCTGTGAAGAACCGCTTTCTTCGCGGTGTCTACCCCTTTCT 60
QY 111 CCCCATATCTGTTTCGACATGAGCTGAGGCGACGCTCGCGGCGGTGAGCCCTGTTCGCA 170
Db 61 CCCCATATCTGTTTCGACATGAGCTGAGGCGACGCTCGCGGCGGTGAGCCCTGTTCGCA 120
QY 171 GCTACGGCGAGAGGGCGCGATTCTTCTTTGTCGCTCCGCTTAGTGGCCGCTCA 230
Db 121 GCTACGGCGAGAGGGCGCGATTCTTCTTTGTCGCTCCGCTTAGTGGCCGCTCA 180
QY 231 TTCCCGCGGTGTCCCGATTATTTAGGGTAGGGAGAGTGTGAGCTTCAGGATCCGAGG 290
Db 181 TTCCCGCGGTGTCCCGATTATTTAGGGTAGGGAGAGTGTGAGCTTCAGGATCCGAGG 240
QY 291 CGTGGCGGCCCATGCGCCCGCTGGGAGCGGCCCGCGGCTGGTACTGCTGTTCAGCGGC 350
Db 241 CGTGGCGGCCCATGCGCCCGCTGGGAGCGGCCCGCGGCTGGTACTGCTGTTCAGCGGC 300
QY 351 RAGAGAAATCCGGAAGACTTCTGACCGAGCGCTGCAGAGCAGACTTGGAGTGTAT 410
Db 301 RAGAGAAATCCGGAAGACTTCTGACCGAGCGCTGCAGAGCAGACTTGGAGTGTAT 360
QY 411 GTCTGTGCTGTCTCTCCGCTCTCTGCTCCACTCAAGGAACAGTATGCTCAGGAGCATGCG 470
Db 361 GTCTGTGCTGTCTCTCCGCTCTCTGCTCCACTCAAGGAACAGTATGCTCAGGAGCATGCG 420
QY 471 TTGAACCTTCAGAGACTCTCTGGACACCGAGCTTCAAGAGAGCCCTTTCGGAAGACATG 530
Db 421 TTGAACCTTCAGAGACTCTCTGGACACCGAGCTTCAAGAGAGCCCTTTCGGAAGACATG 480
QY 531 ATCCGCTGGGAGAGAGAAACCGCAGGCTGACCCAGGCTTCTTTGAGGAAGATTGTG 590
Db 481 ATCCGCTGGGAGAGAGAAACCGCAGGCTGACCCAGGCTTCTTTGAGGAAGATTGTG 540
QY 591 GAGGAGCATCTCCAGCCCATCTGGCTGTGAGTGTACACAGGAGAGTGTCTGACATCCAG 650
Db 541 GAGGAGCATCTCCAGCCCATCTGGCTGTGAGTGTACACAGGAGAGTGTCTGACATCCAG 600
QY 651 TGGTTTCGGGAGGCTATGGGGCCCTGACGAGCGCTCCGCTTGTAGCGTTGGAGCAG 710
Db 601 TGGTTTCGGGAGGCTATGGGGCCCTGACGAGCGCTCCGCTTGTAGCGTTGGAGCAG 660
QY 711 AGCCGAGCAGACGCGGGTGGGTGTTCAGCCAGGGGTGGAGCATGTGAGTCAAGATGT 770
Db 661 AGCCGAGCAGACGCGGGTGGGTGTTCAGCCAGGGGTGGAGCATGTGAGTCAAGATGT 720

```

QY 771 GGCCTGGACACTTCGGGCACTTTGACTGGCTATCGAGAACCATGAGTTGACAGCGC 830  
 DB 721 GGCCTGGACAACTTCGGGCACTTTGACTGGCTATCGAGAACCATGAGTTGACAGCGC 780  
 QY 831 CTGGAGGAGCAGTCTGGAGAACCTGATAGAAATTTATCCGCT-CCAGACTTTAGTCACTAGG 889  
 DB 781 CTGGAGGAGCAGTCTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAAG 840  
 QY 890 TT-CTAGAGTGAAGCTGGGCGCTCTGAGTGGGGTGG 927  
 DB 841 TTCTAGGAAGTGAAGCTGGGCGCTCTGAGTGGGGTGG 880

RESULT 9  
 BM423359 1059 bp mRNA linear EST 29-JAN-2002  
 LOCUS AGENCOURT\_6402244 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5516495  
 5', mRNA sequence.  
 ACCESSION BM423359  
 VERSION BM423359.1 GI:18391584  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1059)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: DCTD/DP  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: ILGM2017 row: c column: 24  
 High quality sequence stop: 681.  
 Location/Qualifiers  
 1..1059  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5516495"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 41"  
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCACAGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

FEATURES  
 source  
 1..1059  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5516495"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 41"  
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCACAGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Query Match 65.0%; Score 833; DB 12; Length 1059;  
 Best Local Similarity 98.1%; Pred. No. 1.2e-187;  
 Matches 874; Conservative 0; Mismatches 12; Indels 5; Gaps 3;  
 QY 303 ATGGCCCCGCTGGAGAGCGCCCGCGCTGGTACTGCTTCAGCGGCAAGAGGAATCC 362  
 DB 1 ATGGCCCCGCTGGAGAGCGCCCGCGCTGGTACTGCTTCAGCGGCAAGAGGAATCC 60  
 QY 363 GGAAGAGCTCTGTGACCGAGCGCTGCGAGCAGACTTGGAGCTGATCTCTGCTGTC 422  
 DB 61 GGAAGAGCTCTGTGACCGAGCGCTGCGAGCAGACTTGGAGCTGATCTCTGCTGTC 120  
 QY 423 CTCGGCTCTCTGTGCTCACTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAACCTCCAG 482

DB 121 CTCGGCTCTCTGTGCTCACTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAACCTCCAG 180  
 QY 483 AGACTCTCTGGACACCAAGACACCTTACAGAGGCGCTTTCGGAAGGACATGATCCGCTGGGGA 542  
 DB 181 AGACTCTCTGGACACCAAGACACCTTACAGAGGCGCTTTCGGAAGGACATGATCCGCTGGGGA 240  
 QY 543 GAGGAGAAACCGCAGGCTGACCCAGGCTTCTTTTCAGGAAAGATTGTGGAGGCACTCTCC 602  
 DB 241 GAGGAGAAACCGCAGGCTGACCCAGGCTTCTTTTCAGGAAAGATTGTGGAGGCACTCTCC 300  
 QY 603 CAGCCCATCTGCTGGTGAAGTACACACGAGAGTGTCTGATCATCCAGTGGTTTCGGGAG 662  
 DB 301 CAGCCCATCTGCTGGTGAAGTACACACGAGAGTGTCTGATCATCCAGTGGTTTCGGGAG 360  
 QY 663 GCCTATGGGCGCTGACGAGAGCGGTCCGCGTTGTAGCGTTGGAGCAGAGCCGACAGCAG 722  
 DB 361 GCCTATGGGCGCTGACGAGAGCGGTCCGCGTTGTAGCGTTGGAGCAGAGCCGACAGCAG 420  
 QY 723 CGGGGCTGGGTTCACGCCAGGCTGGACGATGCTGAGTCAGATGAGTGGCTGGGCTGGACAC 782  
 DB 421 CGGGGCTGGGTTCACGCCAGGCTGGACGATGCTGAGTCAGATGAGTGGCTGGGCTGGACAC 480  
 QY 783 TTCGGGAGCTTTGACTGGGTTCATCGAGAACCATGAGTTGAAACAGCGCTGGAGGAGCAG 842  
 DB 481 TTCGGGAGCTTTGACTGGGTTCATCGAGAACCATGAGTTGAAACAGCGCTGGAGGAGCAG 540  
 QY 843 TTGAGAACTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGG 902  
 DB 541 TTGAGAACTGATAGAAATTTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGG 600  
 QY 903 CTGGGCGCTGCTGAGTGGGGTGGGGTGGCTGACTCTGCAAAATGGGGTGTCCCGGATCC 962  
 DB 601 CTGGGCGCTGCTGAGTGGGGTGGGGTGGCTGACTCTGCAAAATGGGGTGTCCCGGATCC 660  
 QY 963 TGCCCGAGGTGAGGAAACACACAGGCGGGGTCTAGATTTCTGAGGG--GGTTGGTGGATTT 1020  
 DB 661 TGCCCGAGGTGAGGAAACACACAGGCGGGGTCTAGATTTCTGANGNNNGTGTGGTGGATTT 720  
 QY 1021 GGGCAAGGAGGAAACCTCTGGAGACTCTATTTCTCCATGGGAGAGACAGCCTCTCT 1080  
 DB 721 GGGCAAGGAGGAAACCTCTGGAGACTCTATTTCTCCATGGGAGAGACAGCCTCTCT 780  
 QY 1081 TCAGGAGGAGACTTCCAGGCGAAAGAGGGTGTCTTGGTGT--GCTTGAAGCGGAAACC 1138  
 DB 781 TCAGGAGGAGACTTCCAGGCGAAAGAGGGTGTCTTGGTGT--GCTTGAAGCGGAAACC 840  
 QY 1139 CTGCCATATCCCGAGTCCAGTCCCT--CAGCTGTGGCTTCATCC 1188  
 DB 841 CTGCCATATCCCGAGTCCAGTCCCT--CAGCTGTGGCTTCATCC 891

RESULT 10  
 AL527916/c  
 LOCUS AL527916 Homo sapiens 1109 bp mRNA linear EST 23-MAY-2003  
 DEFINITION cDNA clone CS02027YM23 3-PRIME, mRNA sequence.  
 ACCESSION AL527916  
 VERSION AL527916.2 GI:31065767  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1109)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12791409.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

2

Db 763 TTCCAGAGACTCCTGGACACAGCACCACACAGAGAGGCTTTCGGAAGGACATGAGCCG 704  
 QY 537 TGGGAGAGGAGAAACGCGAGCTGACCCAGGCTTCTTTTCAGGAGAAATGTTGGAGGGC 596  
 Db 703 TGGGAGAGGAGAAACGCGAGCTGACCCAGGCTTCTTTTCAGGAGAAATGTTGGAGGGC 645  
 QY 597 ATCTCCAGCCATCTGGCTGGTGGTGGACACAGGAGAGTGTCTGACATCCAGTGGTTT 656  
 Db 644 ATCTCCAGCCATCTGGCTGGTGGTGGACACAGGAGAGTGTCTGACATCCAGTGGTTT 585  
 QY 657 CGGGAGGCTTATGGGGCGCTGACCGACAGAGCTGCGCG-TTGTAGCGTTGGAGAGAGCCG 715  
 Db 584 CGGGAGGCTTANGGGCCGACGACAGAGCTGCGCGTTGTAGCGTTGGAGAGAGCCG 525  
 QY 716 ACAGCAGCGGGCTGGTGTCTACCGCCAGGGGTGGACATCTGAGTCAAGATGGCCCT 775  
 Db 524 ACAGCAGCGGGCTGGTGTCTACCGCCAGGGGTGGACATCTGAGTCAAGATGGCCCT 465  
 QY 776 GGACAACTTCGGGGACTTTGACTGGGTCTATCGAGAACCTGAGTGTGAACAGCGCTTGA 835  
 Db 464 GGACAACTTCGGGGACTTTGACTGGGNCAGAGAACCCAGAGTNGAAGCGCTTGA 405  
 QY 836 GGAGCAGTTGGAGAACCTGTAGAAATTTATCCGCTCCAGACTTTAGTCACTAGTCTTAG 895  
 Db 404 GGAGCAGTTGGAGAACCTGTAGAAATTTATCCGCTCCAGACTTTAGTCACTAGTCTTAG 345  
 QY 896 GAGTCACTGGGGCTGCTGAGTGGGGTGGGGTGAATCTGCAAAATGGGGTGTCCC 955  
 Db 344 GAGTCACTGGGGCTGCTGAGTGGGGTGGGGTGAATCTGCAAAATGGGGTGTCCC 285  
 QY 956 CGGATCTTCGGCGAGTGGAGAACAGACAGGGGGGTCTAGATCTGAGGGGGTGTGGT 1015  
 Db 284 CGGATCTTCGGCGAGTGGAGAACAGACAGGGGGGTCTAGATCTGAGGGGGTGTGGT 225  
 QY 1016 ATATTGGCAAGCAGGAACTCTGGAGACTCTATTTCTCCATGGGAGAGAGCCAT 1075  
 Db 224 ATATTGGCAAGCAGGAACTCTGGAGACTCTATTTCTCCATGGGAGAGAGCCAT 165  
 QY 1076 GCTCTTCAGGAGAGACTCCAGAGGCAAGAGGGGTGTCTTTGGCTGTGCTTTGAAGGGCAA 1135  
 Db 164 GCTCTTCAGGAGAGACTCCAGAGGCAAGAGGGGTGTCTTTGGCTGTGCTTTGAAGGGCAA 105  
 QY 1136 ACCCTGCCATATCCCGAGTCCAGTCCCTGAGCTGGTGGCTTGCATCTGACTGG 1195  
 Db 104 ACCCTGCCATATCCCGAGTCCAGTCCCTGAGCTGGTGGCTTGCATCTGACTGG 45  
 QY 1196 ATGTTCTCAGCCCTTGTCTCGGCAAGAACCCAGAGCTCC 1236  
 Db 44 ATGTTCTCAGCCCTTGTCTCGGCAAGAACCCAGAGTAC 4

RESULT 12  
 BQ675847  
 LOCUS  
 DEFINITION BQ675847 953 bp mRNA linear EST 15-JUL-2002  
 AGENCOURT 8062687 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6211288  
 5', mRNA sequence.

ACCESSION BQ675847  
 VERSION BQ675847.1 GI:21786681  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 953)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov

COMMENT Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2373 row: i column: 17  
 High quality sequence stop: 651.  
 Location/Qualifiers  
 1..953  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6211288"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_102"  
 /note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed  
 by Ling Hong in the laboratory of Gerald M. Rubin  
 (University of California, Berkeley) using ZAP-cDNA  
 synthesis kit (Stratagene) and Superscript II RT (Life  
 Technologies). Note: this is a NIH\_MGC Library."

## FEATURES

source

## ORIGIN

Query Match 64.3%; Score 823.2; DB 13; Length 953;  
 Best Local Similarity 96.4%; Pred. No. 2.5e-185;  
 Matches 863; Conservative 0; Mismatches 29; Indels 3; Gaps 2;

QY 100 CACCTTTTCTCCCATATCTGTTGGACATGAGTGGGACACGCTCGCGGCGGTGAG 159  
 Db 1 CACCTTTTCTCCCATATCTGTTGGACATGAGTGGGACACGCTCGCGGCGGTGAG 60  
 QY 160 CCTGTTCCAGTACGCGAGAGGGGCGCATGTTGTTGTTGTTGTTGTTGTTGTTGTT 219  
 Db 61 CCTGTTCCAGTACGCGAGAGGGGCGCATGTTGTTGTTGTTGTTGTTGTTGTTGTT 120  
 QY 220 GCGCGGTCATTCGCGCGGTGTCGCGATTTAGGGTAGGAGAGTGTGAGTTCAG 279  
 Db 121 GCGCGGTCATTCGCGCGGTGTCGCGATTTAGGGTAGGAGAGTGTGAGTTCAG 180  
 QY 280 GCATCCGAGGGGTGGCGGCCCATGCGGCCCTGCGAGGGCGCCCGCGGTGTTGTC 339  
 Db 181 GCATCCGAGGGGTGGCGGCCCATGCGGCCCTGCGAGGGCGCCCGCGGTGTTGTC 240  
 QY 340 TGTTCAGCGGCAAGAGAAATCGGAGAGACTTGTGACCGAGGGCTGACAGAGCAGAC 399  
 Db 241 TGTTCAGCGGCAAGAGAAATCGGAGAGACTTGTGACCGAGGGCTGACAGAGCAGAC 300  
 QY 400 TTGGAGCTGATGCTGTGCTGCTCCCGGCTCTCTGGTCCACTCAAGGAACAGTATGCTC 459  
 Db 301 TTGGAGCTGATGCTGTGCTGCTCCCGGCTCTCTGGTCCACTCAAGGAACAGTATGCTC 360  
 QY 460 AGAGCATGGCTTGAATCTTCAGAGACTCTCGACACCGACACCTACAGAGGCGCTTC 519  
 Db 361 AGAGCATGGCTTGAATCTTCAGAGACTCTCGACACCGACACCTACAGAGGCGCTTC 420  
 QY 520 GGAAGGACATGATTCGCTGGGAGAGAGAAACCGCAGGCTCACCCAGGCTTCTTTTGA 579  
 Db 421 GGAAGGACATGATTCGCTGGGAGAGAGAAACCGCAGGCTCACCCAGGCTTCTTTTGA 480  
 QY 580 GGAAGATTTGGAGGCGATCTCCAGCCCATCTGGCTGTGTAGTGACACACGAGAGTGT 639  
 Db 481 GGAAGATTTGGAGGCGATCTCCAGCCCATCTGGCTGTGTAGTGACACACGAGAGTGT 540  
 QY 640 CTGACATCCAGTGGTTTCGGAGGCGCTATGGGCGCGTGTGACGAGAGCGTCCGCTTGTAG 699  
 Db 541 CTGACATCCAGTGGTTTCGGAGGCGCTATGGGCGCGTGTGACGAGAGCGTCCGCTTGTAG 600  
 QY 700 CGTTGGAGCAGAGCCGACAGCAGCGGGGTGGGTGTTACGCCAGGGGTGGAGCGATGCTG 759  
 Db 601 CGTTGGAGCAGAGCCGACAGCAGCGGGGTGGGTGTTACGCCAGGGGTGGAGCGATGCTG 660





**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2083 row: j column: 21  
High quality sequence stop: 623.  
Location/Qualifiers  
1. 1045  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5920244"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

**FEATURES**  
**source**

**ORIGIN**

Query Match 63.8%; Score 816.8; DB 13; Length 1045;  
Best Local Similarity 95.5%; Pred. No. 8.5e-184;  
Matches 882; Conservative 0; Mismatches 36; Indels 6; Gaps 4;  
QY 364 GGAAGACTTCGTGACCGGGCTGCAGAGCAGCTTGGAGCTGCTGCTGCTGCC 423  
DB 1 GGAAGACTTCGTGACCGGGCTGCAGAGCAGCTTGGAGCTGCTGCTGCTGCC 60  
QY 424 TCCGGCTCTCTGCTCACTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAACCTCCAGA 483  
DB 61 TCCGGCTCTCTGCTCACTCAAGGAACAGTATGCTCAGGAGCATGGCTTGAACCTCCAGA 120  
QY 484 GACTCTTGACACACAGCCTCAAGAGGCGCTTTCGGAAGACATGATCCGCTGGGAG 543  
DB 121 GACTCTTGACACACAGCCTCAAGAGGCGCTTTCGGAAGACATGATCCGCTGGGAG 180  
QY 544 AGGAGAAACGCGAGCTGACCCAGGCTTCTTTGAGGAAGATTGTGAGGCGCATCTCCC 603  
DB 181 AGGAGAAACGCGAGCTGACCCAGGCTTCTTTGAGGAAGATTGTGAGGCGCATCTCCC 240  
QY 604 AGCCCATCTGGCTGTGATGACACAGGAGAGTGTCTACATCCAGTGGTTTCGGGAGG 663  
DB 241 AGCCCATCTGGCTGTGATGACACAGGAGAGTGTCTACATCCAGTGGTTTCGGGAGG 300  
QY 664 CCTATGGGCGGTGACGAGCGTCCGCTGTGAGCTTGGAGCAGAGCCGACAGCAGC 723  
DB 301 CCTATGGGCGGTGACGAGCGTCCGCTGTGAGCTTGGAGCAGAGCCGACAGCAGC 360  
QY 724 GGGGCTGGGTGTTTACGCCAGGGGTGGAACATGCTGAGTCAAGTGTGGCTTGGACAAT 783  
DB 361 GGGGCTGGGTGTTTACGCCAGGGGTGGAACATGCTGAGTCAAGTGTGGCTTGGACAAT 420  
QY 784 TCGGGGACTTGTGCTGGTCAATCGAACCATGAGCTTCAACAGGCGCTTGGAGGAGCAGT 843  
DB 421 TCGGGGACTTGTGCTGGTCAATCGAACCATGAGCTTCAACAGGCGCTTGGAGGAGCAGT 480  
QY 844 TGGAGAACCTGTATAGATTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGC 903  
DB 481 TGGAGAACCTGTATAGATTATCCGCTCCAGACTTTAGTCACTAGGTTCTAGGAGTGAGC 540

QY 904 TGGGGCTGCTGAGGTGGGGTGGGGCTGACTCTCAAAAATGGGGGTGTCCCCGATCTCT 963  
DB 541 TGGGGCTGCTGAGGTGGGGTGGGGCTGACTCTCAAAAATGGGGGTGTCCCCGATCTCT 600  
QY 964 GCGGAGGTGAGAAACAGACAGGGGGGTCTAGATTCTGAGGGGGTGTGGTGGATATTGGG 1023  
DB 601 GCGGAGGTGAGAAACAGACAGGGGGGTCTAGATTCTGAGGGGGTGTGGTGGATATTGGG 660  
QY 1024 CAGGAGGAAACCTCTGGAGACCTCATTTTCTCCATGGGGAAGACAGCATGCTCTTCA 1083  
DB 661 CAGGAGGAAACCTCTGGAGACCTCATTTTCTCCATGGGGAAGACAGCATGCTCTTCA 720  
QY 1084 GGAGGAGACTCCAAGGGCAAGAGGGGTGTCTTGGCTGTGCTTGAAGGCAAAACCTGCT 1143  
DB 721 GGAGGAGACTCCAAGGGCAAGAGGGGTGTCTTGGCTGTGCTTGAAGGCAAAACCTGCT 780  
QY 1144 ATATCCCGAGTCCAGTCCCTCAGCTGTGCTGCTTGCATCTGCA-TGTTTC 1201  
DB 781 ATATCCCGAGTCCAGTCCCTCAGCTGTGCTGCTTGCATCTGCA-TGTTTC 840  
QY 1202 TCAGCCCCCTT-GTTCGGGCAAGAACCCAGAGCTCCCC--AGTGTGGATATAATAAAC 1257  
DB 841 TCAGCCCCCTTGGTTCTGGGCAAGAACCCAGAGCTCCCCCAGGGTGGGATACCTAATAAC 900  
QY 1258 CTCTTGGAGCAAAAAA 1281  
DB 901 CTCTTGGAGCCCNAAAAAGAAA 924

**RESULT 15**

AL526453 1031 bp mRNA linear EST 23-MAY-2003  
AL526453 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
LOCUS cDNA clone CS0DC014YL24 5-PRIME, mRNA sequence.  
DEFINITION  
AL526453  
AL526453  
AL526453.2 GI:31064311  
EST.

**ACCESSION**

**VERSION**

**KEYWORDS**

**SOURCE**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

On Feb 13, 2001 this sequence version replaced gi:12789946.

Contact: Genoscope - Centre National de Sequencage

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 530.f For

more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC014DF12QP1&cluster=530.f> Contact :

Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :

<http://fulllength.invitrogen.com/InvitrogenCorporation1600>

Paraday Avenue Genoscope sequence ID : CS0DC014DF12QP1.

**FEATURES**

**source**

1. 1031

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DC014YL24"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

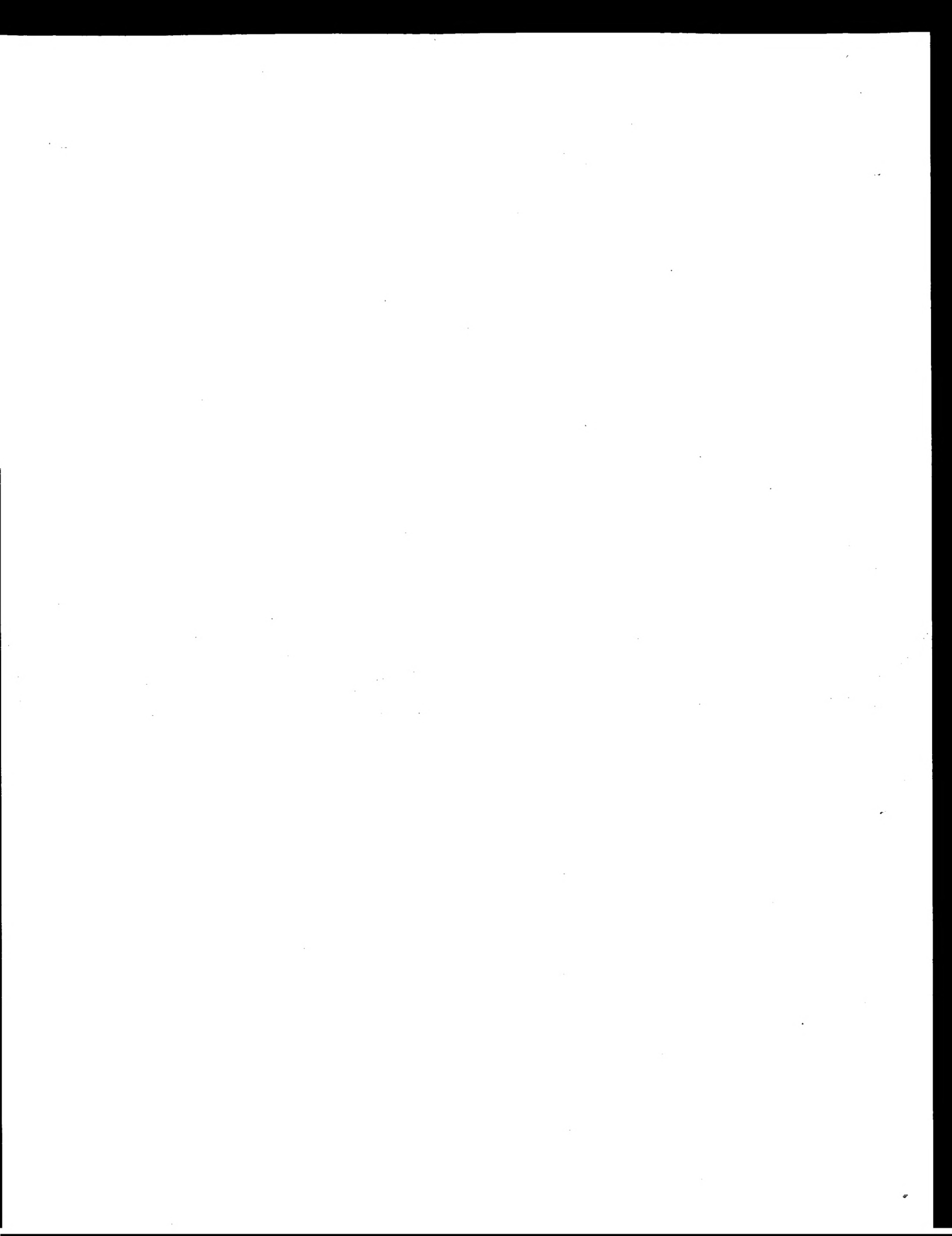
sites of the pCMVSPORT 6 vector. Library was normalized."

**ORIGIN**



Query Match	63.7%	Score 816.4	DB 9	Length 1031	
Best Local Similarity	99.9%	Pred. No. 1.1e-183			
Matches 817	Conservative 0	Mismatches 1	Indels 0	Gaps 0	
Qy	464	GCATGGCTTGAACCTCCAGAGACTCCTGGACACCAAGACCTCTCAAGAGAGCCCTTCGGAA	523		
Db	68	GCATGGCTTGAACCTCCAGAGACTCCTGGACACCAAGACCTCTCAAGAGAGCCCTTCGGAA	127		
Qy	524	GGACATGATCCGCTGGGAGAGAGAAACGCCAGGCTGACCCAGGCTTCCTTTGCAGGAA	583		
Db	128	GGACATGATCCGCTGGGAGAGAGAGAAACGCCAGGCTGACCCAGGCTTCCTTTGCAGGAA	187		
Qy	584	GATTGTGAGGGCATCTCCAGCCCATCTGGCTGGTGAAGTACACACGGAGAGTGTCTGA	643		
Db	188	GATTGTGAGGGCATCTCCAGCCCATCTGGCTGGTGAAGTACACACGGAGAGTGTCTGA	247		
Qy	644	CATCCAGTGGTTTCGGAGAGGCTATGGGGCGTGAACGACACGCTCCGCGTTGTAGCGTT	703		
Db	248	CATCCAGTGGTTTCGGAGAGGCTATGGGGCGTGAACGACACGCTCCGCGTTGTAGCGTT	307		
Qy	704	GGAGCAGAGCCGACAGCAGCGGGCTGGGTGTTCAACCCAGGGGTGACCATCTAGTC	763		
Db	308	GGAGCAGAGCCGACAGCAGCGGGCTGGGTGTTCAACCCAGGGGTGACCATCTAGTC	367		
Qy	764	AGATGTGGCTGACACACTTCGGGGACTTTGACTGGGTGATCGAGAACCATGGAGTTGA	823		
Db	368	AGATGTGGCTGACACACTTCGGGGACTTTGACTGGGTGATCGAGAACCATGGAGTTGA	427		
Qy	824	ACAGCGCTGGAGGAGCAGTTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTC	883		
Db	428	ACAGCGCTGGAGGAGCAGTTGGAGAACCTGATAGAAATTTATCCGCTCCAGACTTTAGTC	487		
Qy	884	ACTAGGTCTTAGAGTGAAGTGGGGGCTGCTGAGGTGGGGTGGGGTGGGGTGGGGTGGGG	943		
Db	488	ACTAGGTCTTAGAGTGAAGTGGGGGCTGCTGAGGTGGGGTGGGGTGGGGTGGGGTGGGG	547		
Qy	944	TGGGGTGTCCCGCATCTCCGCGAGTGAAGAACACAGACAGGGGGGCTCTAGATTCTGA	1003		
Db	548	TGGGGTGTCCCGCGATCTCCGCGAGTGAAGAACACAGACAGGGGGGCTCTAGATTCTGA	607		
Qy	1004	GGGGTGGTGGATATTGGGCAAGGAGGAAACCTCTGGAGACCTCATTTTCTCCATGGG	1063		
Db	608	GGGGTGGTGGATATTGGGCAAGGAGGAAACCTCTGGAGACCTCATTTTCTCCATGGG	667		
Qy	1064	GAAGACAGCCATGCTCTTCAGAGAGAGACTCCAAGGGCAAGAGAGGTGTCTTGGCTGTG	1123		
Db	668	GAAGACAGCCATGCTCTTCAGAGAGAGACTCCAAGGGCAAGAGAGGTGTCTTGGCTGTG	727		
Qy	1124	CTTGAAGCGGAAACCTTGCCATATCCCGAGTCCAGTCCCGCTCAGCTGTGGTGGCTTG	1183		
Db	728	CTTGAAGCGGAAACCTTGCCATATCCCGAGTCCAGTCCCGCTCAGCTGTGGTGGCTTG	787		
Qy	1184	CATCTGACTGGATGTTCTCAGCCCTTGTCTGGCAAGAACCCAGAGCTCCCCAGTGT	1243		
Db	788	CATCTGACTGGATGTTCTCAGCCCTTGTCTGGCAAGAACCCAGAGCTCCCCAGTGT	847		
Qy	1244	GGATACTAATAACCTCTTGGAGCACAAAAA	1281		
Db	848	GGATACTAATAACCTCTTGGAGCACAAAAA	885		

Search completed: March 11, 2004, 14:22:57  
Job time : 2129.06 secs



1	7780.6	99.8	8044	4	US-09-566-921-135	Sequence 135, Appl
2	7423	95.2	7679	4	US-09-220-132-38	Sequence 38, Appl
3	7423	95.2	7680	4	US-09-023-655-1289	Sequence 1289, Ap
4	7423	95.2	7680	5	PCR-US95-09819-6	Sequence 6, Appl
5	7418.2	95.2	7705	1	US-08-259-569-16	Sequence 16, Appl
6	7418.2	95.2	7705	2	US-08-826-885-16	Sequence 16, Appl
7	7413.4	95.1	7705	6	5455158-2	Patent No. 5455158
8	7223	92.7	7803	2	US-08-551-356-1	Sequence 1, Appl
9	7223	92.7	7803	5	PCR-US93-12687-1	Sequence 1, Appl
10	2211.8	28.4	4027	5	US-08-551-356-5	Sequence 5, Appl
11	2211.8	28.4	4027	5	PCR-US93-12687-5	Sequence 5, Appl
12	1400	18.0	2481	4	US-09-366-009-20	Sequence 20, Appl
13	1400	18.0	2481	4	US-08-809-1568-20	Sequence 20, Appl
14	921.2	11.8	1722	4	US-09-366-009-33	Sequence 33, Appl
15	921.2	11.8	1722	4	US-08-809-1568-33	Sequence 33, Appl
16	877.4	11.3	1416	4	US-09-366-009-27	Sequence 27, Appl
17	877.4	11.3	1416	4	US-08-809-1568-27	Sequence 27, Appl
18	860.8	11.0	1644	4	US-09-366-009-17	Sequence 17, Appl
19	860.8	11.0	1644	4	US-08-809-1568-17	Sequence 17, Appl
20	831	10.7	1374	4	US-09-366-009-26	Sequence 26, Appl
21	831	10.7	1374	4	US-08-809-1568-26	Sequence 26, Appl
22	766.4	9.8	986	1	US-07-637-250A-8	Sequence 8, Appl
23	766.4	9.8	986	1	US-08-145-061-8	Sequence 8, Appl
24	652	8.4	653	4	US-09-404-879A-370	Sequence 370, App
25	610	7.8	610	2	US-08-692-787-7	Sequence 7, Appl
26	610	7.8	610	3	US-09-097-199-7	Sequence 7, Appl
27	606	7.8	606	1	US-07-803-623B-10	Sequence 10, Appl

QY 361 GCAGGCTCAGCAAAATGGTTTCAGCCCGCCTCCCGGTGGTGTCTCAGTCAAAAGCAAGCCCGG 420  
 Db 361 GCAGGCTCAGCAAAATGGTTTCAGCCCGCCTCCCGGTGGTGTCTCAGTCAAAAGCAAGCCCGG 420  
 QY 421 TTGTTTATGACAAATGAAACCACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGG 480  
 Db 421 TTGTTTATGACAAATGAAACCACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGG 480  
 QY 481 CAAATCGGTTGGTTTCTACTTGTATGAGAGAGCGGAGGTTTAACTGCGAAAGTAAACC 540  
 Db 481 CAAATCGGTTGGTTTCTACTTGTATGAGAGAGCGGAGGTTTAACTGCGAAAGTAAACC 540  
 QY 541 TGAAGCTGAAGAGACTTCTTTTGAACAATGACACTGGGAACCTTACCGAGTGGGTGACAC 600  
 Db 541 TGAAGCTGAAGAGACTTCTTTTGAACAATGACACTGGGAACCTTACCGAGTGGGTGACAC 600  
 QY 601 TTATGAGCGTCTTAAGACTCCATCATCTGGAGCTGTACTGTGATCGGGGCTGGGCGAGG 660  
 Db 601 TTATGAGCGTCTTAAGACTCCATCATCTGGAGCTGTACTGTGATCGGGGCTGGGCGAGG 660  
 QY 661 GAGAAATGAGCTGTACCATGCGCAAAACCGCTGCGCATGAAGGGGGTCACTCTCAAGATTTG 720  
 Db 661 GAGAAATGAGCTGTACCATGCGCAAAACCGCTGCGCATGAAGGGGGTCACTCTCAAGATTTG 720  
 QY 721 TCACACCTGGAGAGACCACTGAGACTGGTGTACATGTTAGAGTGTGTCTTGG 780  
 Db 721 TCACACCTGGAGAGACCACTGAGACTGGTGTACATGTTAGAGTGTGTCTTGG 780  
 QY 781 TAATGGAAAGAGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGTTGATCATGCTGC 840  
 Db 781 TAATGGAAAGAGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGTTGATCATGCTGC 840  
 QY 841 TGGGACTTCTATGTGGTGGGAAACGTTGGAGAACGCTACCAAGCTGGATGATGGT 900  
 Db 841 TGGGACTTCTATGTGGTGGGAAACGTTGGAGAACGCTACCAAGCTGGATGATGGT 900  
 QY 901 AGATTGCTACTTGGTGGGAGAGGAGCGGACGCTACCTTCTAGAAATAGATG 960  
 Db 901 AGATTGCTACTTGGTGGGAGAGGAGCGGACGCTACCTTCTAGAAATAGATG 960  
 QY 961 CAAAGTACAGGACACAGAGACATCTATAGAAATGGAGACCTGGAGCAAGAGGATAA 1020  
 Db 961 CAAAGTACAGGACACAGAGACATCTATAGAAATGGAGACCTGGAGCAAGAGGATAA 1020  
 QY 1021 TCAGGAAACCTGCTCCAGTGTATCTGACAGGCAACGGCCGAGGAGAGTGAAGTGTGA 1080  
 Db 1021 TCAGGAAACCTGCTCCAGTGTATCTGACAGGCAACGGCCGAGGAGAGTGAAGTGTGA 1080  
 QY 1081 GAGGCAACCTGCTGTGAGAGACCACTGAGCGGATCTGSCCCCTTCAACGATGTTCTGTC 1140  
 Db 1081 GAGGCAACCTGCTGTGAGAGACCACTGAGCGGATCTGSCCCCTTCAACGATGTTCTGTC 1140  
 QY 1141 AGCTGTTTACCAACCGAGGCTCAGCCCGAGGCTCTCTATGCGCACTGTGTACACAGA 1200  
 Db 1141 AGCTGTTTACCAACCGAGGCTCAGCCCGAGGCTCTCTATGCGCACTGTGTACACAGA 1200  
 QY 1201 CAGTGTGTGGTCTACTCTGTGGGATGAGTGGTGGTGAAGACACAGGAAATAGCAAAAT 1260  
 Db 1201 CAGTGTGTGGTCTACTCTGTGGGATGAGTGGTGGTGAAGACACAGGAAATAGCAAAAT 1260  
 QY 1261 GCTTTGCACTGCTGGCAACGGAGTCACTGCGCAAGAGAGAGCTGTAAACCCAGACTTA 1320  
 Db 1261 GCTTTGCACTGCTGGCAACGGAGTCACTGCGCAAGAGAGAGCTGTAAACCCAGACTTA 1320  
 QY 1321 CGGTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGAGAGCGTT 1380  
 Db 1321 CGGTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGAGAGCGTT 1380  
 QY 1381 CTACTCTGACACAGAGAGGCGAGAGCGAGACTCTTTGGTGGAGCACAACCTTCGAA 1440  
 Db 1381 CTACTCTGACACAGAGAGGCGAGAGCGAGACTCTTTGGTGGAGCACAACCTTCGAA 1440

QY 1441 TTATGAGCAGGACCAAGAAATACTCTTTCTGCAAGACCACTCTGTTTGGTTTCAGACTCG 1500  
 Db 1441 TTATGAGCAGGACCAAGAAATACTCTTTCTGCAAGACCACTCTGTTTGGTTTCAGACTCG 1500  
 QY 1501 AGGAGAAATTCCAATGGTGGCTTGTGCCACTTCCCTTCTATACAACCAACAATTA 1560  
 Db 1501 AGGAGAAATTCCAATGGTGGCTTGTGCCACTTCCCTTCTATACAACCAACAATTA 1560  
 QY 1561 CACTGATTCACCTTCTGAGGCGAGAGACAACTAAGTGGTGGTGGGACCAACAGAA 1620  
 Db 1561 CACTGATTCACCTTCTGAGGCGAGAGACAACTAAGTGGTGGTGGGACCAACAGAA 1620  
 QY 1621 CTATGATGCGGACCAAGAGTGTGGGTTCTGCCCTGCTGCCACGAGGAAATCTGCAC 1680  
 Db 1621 CTATGATGCGGACCAAGAGTGTGGGTTCTGCCCTGCTGCCACGAGGAAATCTGCAC 1680  
 QY 1681 AACCAATGAAGGGGTCATGTACCGCATTTGGAGATCAGTGGGATAAGACGATGACATGG 1740  
 Db 1681 AACCAATGAAGGGGTCATGTACCGCATTTGGAGATCAGTGGGATAAGACGATGACATGG 1740  
 QY 1741 TCACATGATGAGGTGCACTGTGTGGGAAATGGTGGTGGGAAATGGACATGATTCSCCTA 1800  
 Db 1741 TCACATGATGAGGTGCACTGTGTGGGAAATGGTGGTGGGAAATGGACATGATTCSCCTA 1800  
 QY 1801 CTCGAGCTTCGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACAT 1860  
 Db 1801 CTCGAGCTTCGAGATCAGTGCATTTGATGACATCACTTACAATGTGAACGACACAT 1860  
 QY 1861 CCAAGAGCTCATGAGAGGGGACATGCTGACATGATGACATCACTTACAATGTGAACGACACAT 1920  
 Db 1861 CCAAGAGCTCATGAGAGGGGACATGCTGACATGATGACATCACTTACAATGTGAACGACACAT 1920  
 QY 1921 CAGGTGGAAAGTGTGATCCCGTGCAGCAATCCAGGATTCAGAGACTGGGACGTTTTATCA 1980  
 Db 1921 CAGGTGGAAAGTGTGATCCCGTGCAGCAATCCAGGATTCAGAGACTGGGACGTTTTATCA 1980  
 QY 1981 AATTGAGATTCATGGAGAGATGATGATGGTGTGATGATGATGATGATGATGATGATG 2040  
 Db 1981 AATTGAGATTCATGGAGAGATGATGATGGTGTGATGATGATGATGATGATGATGATG 2040  
 QY 2041 CCGTGGCATTTGGGAGTGGCATTTGCCAACTTTACAGACCTATCCAAAGCTCAAGTGGTCC 2100  
 Db 2041 CCGTGGCATTTGGGAGTGGCATTTGCCAACTTTACAGACCTATCCAAAGCTCAAGTGGTCC 2100  
 QY 2101 TGTGCAAGTATTTATCACTGAGACTCCGAGTCCAGCCAACTCCACCCCATCCAGTGGAA 2160  
 Db 2101 TGTGCAAGTATTTATCACTGAGACTCCGAGTCCAGCCAACTCCACCCCATCCAGTGGAA 2160  
 QY 2161 TGCAACCAAGCTATCTCACTTTCCAAAGTACATTTCTCAGTGGAGACCTAAATAATCTGT 2220  
 Db 2161 TGCAACCAAGCTATCTCACTTTCCAAAGTACATTTCTCAGTGGAGACCTAAATAATCTGT 2220  
 QY 2221 AGCCCGTTGGAAGAAAGTACCATACAGGCCACTTAAACTCTCTACACCATCAAGGCGCT 2280  
 Db 2221 AGCCCGTTGGAAGAAAGTACCATACAGGCCACTTAAACTCTCTACACCATCAAGGCGCT 2280  
 QY 2281 GAGGCTGTGTGATACAGGGGCGAGCTCATCAGCATCCAGAGTACAGGCGGACCAAGA 2340  
 Db 2281 GAGGCTGTGTGATACAGGGGCGAGCTCATCAGCATCCAGAGTACAGGCGGACCAAGA 2340  
 QY 2341 AGTGACTCGCTTTGACTTACCAACCAAGAGACCAAGACCACTGTGACCAAGCAACCGCT 2400  
 Db 2341 AGTGACTCGCTTTGACTTACCAACCAAGAGACCAAGACCACTGTGACCAAGCAACCGCT 2400  
 QY 2401 GACAGGAGAGACGACTCCCTTTCTCTCTGTTGGGCGACTTCTGAAATCTGTGACCGAAAT 2460  
 Db 2401 GACAGGAGAGACGACTCCCTTTCTCTCTGTTGGGCGACTTCTGAAATCTGTGACCGAAAT 2460  
 QY 2461 CACAGCCAGTAGCTTTGTGTCTCTGGGTCTCAGCTTCGACACCCGTGTGGGATTCGG 2520  
 Db 2461 CACAGCCAGTAGCTTTGTGTCTCTGGGTCTCAGCTTCGACACCCGTGTGGGATTCGG 2520  
 QY 2521 GGTGGAATATGAGCTGAGTGGAGGGAGATGAGCCACAGTACTCTGGATCTTCCAGCAC 2580

2521 GGTGGAATATGAGCTGAGTGAGGAGAGATGAGCCACAGTACCTGGATCTTCCAGCAC 2580  
2581 AGCCACTTCTGTGAAATCATCCCTGACCTGCTTCTGGCCGAAATATCATTTGTAATGTCTA 2640  
2581 AGCCACTTCTGTGAAATCATCCCTGACCTGCTTCTGGCCGAAATATCATTTGTAATGTCTA 2640  
2641 TCAGATATCTGAGATGAGGAGAGAGATTTGATTCCTGTCTACCTTCTCAAAACAGAGGCC 2700  
2641 TCAGATATCTGAGATGAGGAGAGAGATTTGATTCCTGTCTACCTTCTCAAAACAGAGGCC 2700  
2701 TGATGCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760  
2701 TGATGCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760  
2761 GAGCAGACCCAGGCTCCCATCAAGGATACAGAAATAGTCTATTTCGCCATCAGTGAAGG 2820  
2821 TAGCAGCAGAGACTCAACCTTCTGAACTGCAACTCCGTCAGTCCCTGAGTCTGCA 2880  
2821 TAGCAGCAGAGACTCAACCTTCTGAACTGCAACTCCGTCAGTCCCTGAGTCTGCA 2880  
2881 ACCTGGTGTTCAGTATAACATCACTATCTATGCTGTGGAGAAATCAAGAAAGTACACC 2940  
2881 ACCTGGTGTTCAGTATAACATCACTATCTATGCTGTGGAGAAATCAAGAAAGTACACC 2940  
2941 TGTGTCTATTCAGAGAAACCACTGGCACCACCGTCCAGTACAGTGCCTCTCCAG 3000  
2941 TGTGTCTATTCAGAGAAACCACTGGCACCACCGTCCAGTACAGTGCCTCTCCAG 3000  
3001 GGACCTGCACTTGTGGAAGTGACAGACGTGGAAGTCAACATCATGTGGACACCGCTGA 3060  
3001 GGACCTGCACTTGTGGAAGTGACAGACGTGGAAGTCAACATCATGTGGACACCGCTGA 3060  
3061 GAGTGCACTGACCGGCTACCGTGTGAGTATGATTCCTGAGTCCCTGAGGAGCAGCG 3120  
3061 GAGTGCACTGACCGGCTACCGTGTGAGTATGATTCCTGAGTCCCTGAGGAGCAGCG 3120  
3121 GCAGAGCTCCCATCAGCAGGAAACCTTTGAGAGAGTCAACCTGCTGAGTCCCTGAGG 3180  
3121 GCAGAGCTCCCATCAGCAGGAAACCTTTGAGAGAGTCAACCTGCTGAGTCCCTGAGG 3180  
3181 CACTATTACTTCAAACTCTTTCAGTGTGATGATTCCTGAGTCCCTGAGGAGCAGCG 3240  
3181 CACTATTACTTCAAACTCTTTCAGTGTGATGATTCCTGAGTCCCTGAGGAGCAGCG 3240  
3241 TCACAGACACCAACTGAGTGTCTCCCACTAACCTCCAGTTCCTGATGAACTGATTC 3300  
3241 TCACAGACACCAACTGAGTGTCTCCCACTAACCTCCAGTTCCTGATGAACTGATTC 3300  
3301 TACTGTCTCTGAGATGAGTCTCCACTCGGCCCCAGATAACAGGATACCGACTGACCGT 3360  
3301 TACTGTCTCTGAGATGAGTCTCCACTCGGCCCCAGATAACAGGATACCGACTGACCGT 3360  
3361 GGGCTTACCCGAGAGGACAGCCAGGAGTCAATGTGGTCCCTGCTGCTCCAGTA 3420  
3361 GGGCTTACCCGAGAGGACAGCCAGGAGTCAATGTGGTCCCTGCTGCTCCAGTA 3420  
3421 CCCACTGAGGAATCTGAGGCTGATCTGAGTACACCGTATCCCTGCTGAGGAGTAAAGG 3480  
3421 CCCCTGAGGAATCTGAGGCTGATCTGAGTACACCGTATCCCTGCTGAGGAGTAAAGG 3480  
3481 CAACCAAGAGGCCCAAGGCTGAGTCTTACCACTGAGGCTGGAGCTCTAT 3540  
3481 CAACCAAGAGGCCCAAGGCTGAGTCTTACCACTGAGGCTGGAGCTCTAT 3540  
3541 TCCACCTTACCAACCGAGTGTGAGACCACTTGTGATCAGTGGACGCTGCTCC 3600  
3541 TCCACCTTACCAACCGAGTGTGAGACCACTTGTGATCAGTGGACGCTGCTCC 3600  
3601 AAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAGT 3660

3601 AAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAGT 3660  
3661 GACTTCAGACTCAGGAAAGCATCGTTGTGTCGGCTTCACTCCAGGAGTAGAATACGCTTA 3720  
3661 GACTTCAGACTCAGGAAAGCATCGTTGTGTCGGCTTCACTCCAGGAGTAGAATACGCTTA 3720  
3721 CACCATCCAGTCTGAGAGATGACAGGAAAGATGCGCCAAATTTGTAACAAAGTGGT 3780  
3721 CACCATCCAGTCTGAGAGATGACAGGAAAGATGCGCCAAATTTGTAACAAAGTGGT 3780  
3781 GACACCAATTTGCTCCACCAACAACTTTCGATCTGGAGGCAAAACCTGACACTGGAGTGT 3840  
3781 GACACCAATTTGCTCCACCAACAACTTTCGATCTGGAGGCAAAACCTGACACTGGAGTGT 3840  
3841 CACAGTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATTTACCAAC 3900  
3841 CACAGTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATTTACCAAC 3900  
3901 CCTTACAAAACCGCCAGCAGGGAATTTCTTGGAAAGAGTGTCTCATCTGATCAGAGCTC 3960  
3901 CCTTACAAAACCGCCAGCAGGGAATTTCTTGGAAAGAGTGTCTCATCTGATCAGAGCTC 3960  
3961 CTGCACTTTTGAATTAACCTGAGTCCCGCTGGAGTCAAAATGTGAGTGTTCACATGTCAA 4020  
3961 CTGCACTTTTGAATTAACCTGAGTCCCGCTGGAGTCAAAATGTGAGTGTTCACATGTCAA 4020  
4021 GGATGACAAGGAAAGTGTCCCTATCTGTATACCATCATCCAGCTGTTCCTCTCCAC 4080  
4021 GGATGACAAGGAAAGTGTCCCTATCTGTATACCATCATCCAGCTGTTCCTCTCCAC 4080  
4081 TGAAGTGTGATTAACCAACATTTGTCAGACACCAATGCTGTCACCTGGGCTCCACCCCC 4140  
4081 TGAAGTGTGATTAACCAACATTTGTCAGACACCAATGCTGTCACCTGGGCTCCACCCCC 4140  
4141 ATCCATGATTTAAACCACTTCTGTCGTCCTTACTCACCTGTGAAAAATGAGGAAGTGT 4200  
4141 ATCCATGATTTAAACCACTTCTGTCGTCCTTACTCACCTGTGAAAAATGAGGAAGTGT 4200  
4201 TGCAGAGTGTCAATTTCTCCTTCAGACAAATGTCAGTGTCTTAAACAAATCTCTGCTGG 4260  
4201 TGCAGAGTGTCAATTTCTCCTTCAGACAAATGTCAGTGTCTTAAACAAATCTCTGCTGG 4260  
4261 TACAGATATCTGAGTGTCTCCAGTGTCTAGAACCAACATGAGAGACACCTCTTAG 4320  
4261 TACAGATATCTGAGTGTCTCCAGTGTCTAGAACCAACATGAGAGACACCTCTTAG 4320  
4321 AGGAAGACAGAAAAACAGTCTTGAATTCCTCCAACTGGCAATTTCTGATATTACTGC 4380  
4321 AGGAAGACAGAAAAACAGTCTTGAATTCCTCCAACTGGCAATTTCTGATATTACTGC 4380  
4381 CAACTCTTTTACTGTGCACTGGATGTCTTCGAGGCCACCATCACTGGCTACAGATCCG 4440  
4381 CAACTCTTTTACTGTGCACTGGATGTCTTCGAGGCCACCATCACTGGCTACAGATCCG 4440  
4441 CCATCATCCCGAGCATTTCAGTGGAGACCTTCGAGAGATCGGGTGGCCCACTCTCGAA 4500  
4441 CCATCATCCCGAGCATTTCAGTGGAGACCTTCGAGAGATCGGGTGGCCCACTCTCGAA 4500  
4501 TTCATCATCCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAGCATCGTGTCT 4560  
4501 TTCATCATCCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAGCATCGTGTCT 4560  
4561 TAATGGCAGAGAGGAAGTCCCTTATTGATTTGGCCAAACAAATCAACAGTTCCTGATTTCC 4620  
4561 TAATGGCAGAGAGGAAGTCCCTTATTGATTTGGCCAAACAAATCAACAGTTCCTGATTTCC 4620  
4621 GAGGACCTGGAAGTGTGTGCGACCCCAACAGCTCTGATTCAGTGGGATGCTCC 4680  
4621 GAGGACCTGGAAGTGTGTGCGACCCCAACAGCTCTGATTCAGTGGGATGCTCC 4680  
4681 TGCTGTCACTGAGATATTACAGGATCACTTACGAGAAAACAGAGGAAATAGCCCTGT 4740  
4681 TGCTGTCACTGAGATATTACAGGATCACTTACGAGAAAACAGAGGAAATAGCCCTGT 4740







QY 1249 AAATAAGCAATGCTTTGACAGTGCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT 1308  
Db 901 AAATAAGCAATGCTTTGACAGTGCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT 960  
QY 1309 AACCCAGACTTACGGTGCACACTCAATGGAGAGCCATGTCTTACCATTCACCTACAA 1368  
Db 961 AACCCAGACTTACGGTGCACACTTAAATGGAGAGCCATGTCTTACCATTCACCTACAA 1020  
QY 1369 TGGCAGGACGTTTACTCTCTGCACACAGAAAGGCGACAGACCGGACATCTTTGGTGCAG 1428  
Db 1021 TGGCAGGACGTTTACTCTCTGCACACAGAAAGGCGACAGACCGGACATCTTTGGTGCAG 1080  
QY 1429 CACAACTTGGAAATATGAGCAGGACAGAAATATCTTTCTTGGCAGACACACCTGTTTT 1488  
Db 1081 CACAACTTGGAAATATGAGCAGGACAGAAATATCTTTCTTGGCAGACACACCTGTTTT 1140  
QY 1489 GGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCCATACAA 1548  
Db 1141 GGTTCAGACTCAAGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCCATACAA 1200  
QY 1549 CAAACCAATTAACATGATGACCTTCTGAGGGGAGAGACAAACATGAAGTGGTGG 1608  
Db 1201 CAAACCAATTAACATGATGACCTTCTGAGGGGAGAGACAAACATGAAGTGGTGG 1260  
QY 1609 GACCACAGAACTATGATGCCAGACAGAAAGTTTGGGTTCTGCCCATGGCTGCCACGA 1668  
Db 1261 GACCACAGAACTATGATGCCAGACAGAAAGTTTGGGTTCTGCCCATGGCTGCCACGA 1320  
QY 1669 GGAATCTGCAACCAATGAAGGGGTCAATGACCGCATTTGGAGATCAAGTGGGTAAGCA 1728  
Db 1321 GGAATCTGCAACCAATGAAGGGGTCAATGACCGCATTTGGAGATCAAGTGGGTAAGCA 1380  
QY 1729 GCATGACATGGGTCAATGATGACGATGACGCTGTCTGGGAATGCTGTGGGATGGAC 1788  
Db 1381 GCATGACATGGGTCAATGATGACGATGACGCTGTCTGGGAATGCTGTGGGATGGAC 1440  
QY 1789 ATGCATTTGCTTACTCGCAGCTTCGAGATCAGTGCATTTGTGATGACATCACTTCAATGT 1848  
Db 1441 ATGCATTTGCTTACTCGCAGCTTCGAGATCAGTGCATTTGTGATGACATCACTTCAATGT 1500  
QY 1849 GAAAGCACAATTCACAGCTTCAGAGAGGGGACATGCTGACATCTACATGCTTCGG 1908  
Db 1501 GAAAGCACAATTCACAGCTTCAGAGAGGGGACATGCTGACATCTACATGCTTCGG 1560  
QY 1909 TCAGGCTCGGGGACAGTGAAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968  
Db 1561 TCAGGCTCGGGGACAGTGAAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620  
QY 1969 GACGTTTTATCAAAATGGAGATTCATGGAGAGATGTCATGCTGTGATGATCAGTGG 2028  
Db 1621 GACGTTTTATCAAAATGGAGATTCATGGAGAGATGTCATGCTGTGATGATCAGTGG 1680  
QY 2029 CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATGCCAACCTTTACAGACCTATCCCAAG 2088  
Db 1681 CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATGCCAACCTTTACAGACCTATCCCAAG 1740  
QY 2089 CTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGATCGAGTCCCACTCCACCC 2148  
Db 1741 CTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGATCGAGTCCCACTCCCACTCC 1800  
QY 2149 CATCCAGTGGAAATGACCAAGCCATCTCACATTTCCAGTATCTTCTCAGGTGGAGACC 2208  
Db 1801 CATCCAGTGGAAATGACCAAGCCATCTCACATTTCCAGTATCTTCTCAGGTGGAGACC 1860  
QY 2209 TAAAAATCTGTAGGCGGTGGAGAAAGCTACCATPACAGGCGCATTTAACTCCTACAC 2268  
Db 1861 TAAAAATCTGTAGGCGGTGGAGAAAGCTACCATPACAGGCGCATTTAACTCCTACAC 1920  
QY 2269 CATCAAGGCTCAAGCCTGTGTATATACAGGGCCAGCTCATCAGCATCCAGCTA 2328  
Db 1921 CATCAAGGCTCAAGCCTGTGTATATACAGGGCCAGCTCATCAGCATCCAGCTA 1980

QY 2329 CGGCCACCAAGAGTGAATGCTCGCTTGAATCTTCAACACCAAGAGAGATGAGCCACAGCTGTGAC 2388  
Db 1981 CGGCCACCAAGAGTGAATGCTCGCTTGAATCTTCAACACCAAGAGAGATGAGCCACAGCTGTGAC 2040  
QY 2389 CAGCAACACCGTGCAGAGAGAGACGACTCCCTTTTCTCTCTTTGTGGCCACTTCTTGAATC 2448  
Db 2041 CAGCAACACCGTGCAGAGAGAGACGACTCCCTTTTCTCTCTTTGTGGCCACTTCTTGAATC 2100  
QY 2449 TGTGACCGAAATACAGCCAGTAGCTTTGTGGTCTCTCTGGTCTCAGCTTCGACACCGT 2508  
Db 2101 TGTGACCGAAATACAGCCAGTAGCTTTGTGGTCTCTCTGGTCTCAGCTTCGACACCGT 2160  
QY 2509 GTCCGGATTTCCGGTGGAAATATGAGCTGAGTGAGGAGAGATGAGCCACAGTACTGGA 2568  
Db 2161 GTCCGGATTTCCGGTGGAAATATGAGCTGAGTGAGGAGAGATGAGCCACAGTACTGGA 2220  
QY 2569 TCTTCCAGACAGCCACTTCTGTGAAACATCCCTGACCTGCTTCTGGCCGAAATACAT 2628  
Db 2221 TCTTCCAGACAGCCACTTCTGTGAAACATCCCTGACCTGCTTCTGGCCGAAATACAT 2280  
QY 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTTACTTCA 2688  
Db 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTTACTTCA 2340  
QY 2689 AACCAACAGCCCTGATGCTCCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2748  
Db 2341 AACCAACAGCCCTGATGCTCCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2400  
QY 2749 TGTGTTTCTGCTGGAGCAGACCCAGCTCCATCAAGGGTACAGATAGTCTATTTCGGC 2808  
Db 2401 TGTGTTTCTGCTGGAGCAGACCCAGCTCCATCAAGGGTACAGATAGTCTATTTCGGC 2460  
QY 2809 ATCAGTAGAAGTGTAGCAGACAGAACTCAACCTTCTGAAAATGCAAACTCCGTCACT 2868  
Db 2461 ATCAGTAGAAGTGTAGCAGACAGAACTCAACCTTCTGAAAATGCAAACTCCGTCACT 2520  
QY 2869 CAGTGTGCTGCAACCTGCTGTTTCAATTAACATCACTATCTATGCTGTGAGAAATCA 2928  
Db 2521 CAGTGTGCTGCAACCTGCTGTTTCAATTAACATCACTATCTATGCTGTGAGAAATCA 2580  
QY 2929 AGAAAGTACACCTGTTTGTCAATTCACCAAGAAACCACTGGCACCACCGTTCAGATACAGT 2988  
Db 2581 AGAAAGTACACCTGTTTGTCAATTCACCAAGAAACCACTGGCACCACCGTTCAGATACAGT 2640  
QY 2989 GCCTCTCCAGGGACCTGCAAGTTTGTGAAATGACAGACGTTGAAAGTCAACATCATGTG 3048  
Db 2641 GCCTCTCCAGGGACCTGCAAGTTTGTGAAATGACAGACGTTGAAAGTCAACATCATGTG 2700  
QY 3049 GACACCGCTGAGAGTGCAGTGCAGCCGCTACCGTGTGGATGTGATCCCCGTCAACCTGCC 3108  
Db 2701 GACACCGCTGAGAGTGCAGTGCAGCCGCTACCGTGTGGATGTGATCCCCGTCAACCTGCC 2760  
QY 3109 TGGCGAGCAGCGGCAGAGGCTGCCATCAGCAGGAAACCTTTGAGAGATCAGCGGCT 3168  
Db 2761 TGGCGAGCAGCGGCAGAGGCTGCCATCAGCAGGAAACCTTTGAGAGATCAGCGGCT 2820  
QY 3169 GTCCCTTGGGCTCACCTTATTTCTTCAAGTCTTTGAGTGAAGTGGAGGAGAGAGCAA 3228  
Db 2821 GTCCCTTGGGCTCACCTTATTTCTTCAAGTCTTTGAGTGAAGTGGAGGAGAGCAA 2880  
QY 3229 GCCTCTGACTGTCTAACAGACAAACCTGAGTGTGATGCTCCCACTCAACCTGTCTCAA 3288  
Db 2881 GCCTCTGACTGTCTAACAGACAAACCTGAGTGTGATGCTCCCACTCAACCTGTCTCAA 2940  
QY 3289 TGAACCTGATTTCTTACTGCTGTGATGAGTGTGACTCCAGCTCGGGCCAGATACAGGATA 3348  
Db 2941 TGAACCTGATTTCTTACTGCTGTGATGAGTGTGACTCCAGCTCGGGCCAGATACAGGATA 3000  
QY 3349 CCGACTGACCGTGGGCTTACCCGAAAGAGACAGCCAGGACGTAACAATGTGGTCCCTC 3408  
Db 3001 CCGACTGACCGTGGGCTTACCCGAAAGAGACAGCCAGGACGTAACAATGTGGTCCCTC 3060  
QY 3409 TGTCTCCAAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTCGT 3468

Db 3061 TGTCTCAAGTACCCCTGAGGATCTGCAGCTGCATCTGATACACCGTATCCCTCGT 3120  
2y 3469 GGCATTAAGGGCAACCAAGAGAGCCCAAGCACTGGAGTCTTTACACACCTGCAGCC 3528  
Db 3121 GGCATTAAGGGCAACCAAGAGAGCCCAAGCACTGGAGTCTTTACACACCTGCAGCC 3180  
2y 3529 TGGGAGTCTTATTCACCTTACACACCGAGGTGACTGAGACCAACATTTGTGATCACATG 3588  
Db 3181 TGGGAGTCTTATTCACCTTACACACCGAGGTGACTGAGACCAACATTTGTGATCACATG 3240  
2y 3589 GACCCCTGCTCCAAAGATTGGTTTAAGCTGGGTGATACGACCAAGCCAGGAGGAGCC 3648  
Db 3241 GACCCCTGCTCCAAAGATTGGTTTAAGCTGGGTGATACGACCAAGCCAGGAGGAGCC 3300  
2y 3649 ACCACGAGAGTGTACTTACAGCTCAGAAAGCTGTTGTTGCTCCGGCTTGACTCCAGAGT 3708  
Db 3301 ACCACGAGAGTGTACTTACAGCTCAGAAAGCTGTTGTTGCTCCGGCTTGACTCCAGAGT 3360  
2y 3709 AGAATAGTCTTACACCATCCAAAGTCTTGAGAGATGGACAGAAAGAGATGCGCCAAATTGT 3768  
Db 3361 AGAATAGTCTTACACCATCCAAAGTCTTGAGAGATGGACAGAAAGAGATGCGCCAAATTGT 3420  
2y 3769 AAACAAAGTGTGACACATTTGTCACCAACAAACCTTGCATCTGGAGGCAACCTCGA 3828  
Db 3421 AAACAAAGTGTGACACATTTGTCACCAACAAACCTTGCATCTGGAGGCAACCTCGA 3480  
2y 3829 CACTGGAGTGTCTCAGTCTCTGGAGAGGAGACCAACCCACAGACATTAAGTGGTTATAG 3888  
Db 3481 CACTGGAGTGTCTCAGTCTCTGGAGAGGAGACCAACCCACAGACATTAAGTGGTTATAG 3540  
2y 3889 AATTACCAACACCCCTACAAACGGCCAGAGGAGGAAATCTTTGGAAGAGTGTCCATGC 3948  
Db 3541 AATTACCAACACCCCTACAAACGGCCAGAGGAGGAAATCTTTGGAAGAGTGTCCATGC 3600  
2y 3949 TGATCAGAGTCTCTGCACTTTGATACCTGAGTCCCGGCTTGAGTACAAATGTCAAGTGT 4008  
Db 3601 TGATCAGAGTCTCTGCACTTTGATACCTGAGTCCCGGCTTGAGTACAAATGTCAAGTGT 3660  
2y 4009 TTACACTGTGAGAGTACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
Db 3661 TTACACTGTGAGAGTACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
2y 4069 TCCCTCTCCACACTGACCTGGATTCACCAACATTTGTCAGACACCATGGTGTCACTCTG 4128  
Db 3721 TCCCTCTCCACACTGACCTGGATTCACCAACATTTGTCAGACACCATGGTGTCACTCTG 3780  
2y 4129 GGCTCCACCCCCATCCATTTGATTTAACCAACTTCTCTGGTGGTGTACTCACTGTGAAAA 4188  
Db 3781 GGCTCCACCCCCATCCATTTGATTTAACCAACTTCTCTGGTGGTGTACTCACTGTGAAAA 3840  
2y 4189 TGAGGAAGATGTTGCAGAGTGTCAATTTCTCTTTCAGACAAATGCGAGTGTCTTAAACAA 4248  
Db 3841 TGAGGAAGATGTTGCAGAGTGTCAATTTCTCTTTCAGACAAATGCGAGTGTCTTAAACAA 3900  
2y 4249 TCTCTGCTGGTACAGAAATGATGAGTGTCTCAGTGTCTCAGTGTCTCAGAAACAAATGAGAG 4308  
Db 3901 TCTCTGCTGGTACAGAAATGATGAGTGTCTCAGTGTCTCAGTGTCTCAGAAACAAATGAGAG 3960  
2y 4309 CACACCTTTAGAGAGACAGAAACAGGTCTTGAATTCCTCCCACTGGCACTTGCATTTTC 4368  
Db 3961 CACACCTTTAGAGAGACAGAAACAGGTCTTGAATTCCTCCCACTGGCACTTGCATTTTC 4020  
2y 4369 TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCACTGG 4428  
Db 4021 TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCACTGG 4080  
2y 4429 CTACAGGATCCGCCATCATCCGAGGCACTTCAAGTGGAGAGCCTCGAGAAAGATCGGTGCC 4488  
Db 4081 CTACAGGATCCGCCATCATCCGAGGCACTTCAAGTGGAGAGCCTCGAGAAAGATCGGTGCC 4140  
2y 4489 CCACTCTCGGAATTCATCAACCTCACCCTCAGGACATCTCCAGGACAGAGTATGTGTGTCAG 4548

Db 4141 CCACCTCTCGGAATTCATCACCTCACCAACCTCACTCCAGGCACAGAGTATGTGGTTCAG 4200  
Qy 4549 CATCGTCTCTTAAATGCGACAGAGGAAGTCCCTTATTGATGGCCCAACATCAACAGT 4608  
Db 4201 CATCGTCTCTTAAATGCGACAGAGGAAGTCCCTTATTGATGGCCCAACATCAACAGT 4260  
Qy 4609 TTCTGATGTTCCGAGGACCTGGAAGTTGTTGCTGCGACCCCAACCTCACTGATCAG 4668  
Db 4261 TTCTGATGTTCCGAGGACCTGGAAGTTGTTGCTGCGACCCCAACCTCACTGATCAG 4320  
Qy 4669 CTGGGATGCTCTGCTGTACAGTGAGATATTAAGATCACTTACGAGAGAAACAGGAGG 4728  
Db 4321 CTGGGATGCTCTGCTGTACAGTGAGATATTAAGATCACTTACGAGAGAAACAGGAGG 4380  
Qy 4729 AAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGTACCATCAGCGG 4788  
Db 4381 AAATAGCCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGTACCATCAGCGG 4440  
Qy 4789 COTTAACCTGGAGTGTGATTAACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4848  
Db 4441 COTTAACCTGGAGTGTGATTAACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4500  
Qy 4849 CCCCACAGCAGCAAGCCAAATTTCCATTAATTAACGAAACAGAAATTCACAAACCATCCCA 4908  
Db 4501 CCCCACAGCAGCAAGCCAAATTTCCATTAATTAACGAAACAGAAATTCACAAACCATCCCA 4560  
Qy 4909 GATCAAGTGAACGATGTTACAGCAACAGCATTTAGTGTCAAGTGGTGCCTTCAAGTTC 4968  
Db 4561 GATCAAGTGAACGATGTTACAGCAACAGCATTTAGTGTCAAGTGGTGCCTTCAAGTTC 4620  
Qy 4969 CCGCTTACTGTTACAGAGTAAACCCACCTCCCAAAATGGAAGGAGGAGGACCAACAAAC 5028  
Db 4621 CCGCTTACTGTTACAGAGTAAACCCACCTCCCAAAATGGAAGGAGGAGGACCAACAAAC 4680  
Qy 5029 TAAACTCCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACACAGTGA 5088  
Db 4681 TAAACTCCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACACAGTGA 4740  
Qy 5089 GTATGTTGTTAGTGTCTAGTCTAGATCCAAAGGAGGAGGAGTCACTCTGTTTCAGAC 5148  
Db 4741 GTATGTTGTTAGTGTCTAGTCTAGATCCAAAGGAGGAGGAGTCACTCTGTTTCAGAC 4800  
Qy 5149 TGCAGTACCAACATTTGATCGCCCTTAAAGGACTGGCATTTCACTGATGTGATGTCGATTC 5208  
Db 4801 TGCAGTACCAACATTTGATCGCCCTTAAAGGACTGGCATTTCACTGATGTGATGTCGATTC 4860  
Qy 5209 CATCAAAATTCGTTGGGAAAGCCCAAGGAGGAGGAGTTCAGGTCAGAGGTGACCTACTC 5268  
Db 4861 CATCAAAATTCGTTGGGAAAGCCCAAGGAGGAGGAGTTCAGGTCAGAGGTGACCTACTC 4920  
Qy 5269 GAGCCCTCAGAGTGGAAATCCCATGAGCTATTCCTCTGCACTGATGTTGAAGAGACACTGC 5328  
Db 4921 GAGCCCTCAGAGTGGAAATCCCATGAGCTATTCCTCTGCACTGATGTTGAAGAGACACTGC 4980  
Qy 5329 AGAGTGCAGAGCTCAGACCGGTTCTGAGTACAGTCAAGTGTGTTGCTTGCACGA 5388  
Db 4981 AGAGTGCAGAGCTCAGACCGGTTCTGAGTACAGTCAAGTGTGTTGCTTGCACGA 5040  
Qy 5389 TGATATGAGAGGACGAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5448  
Db 5041 TGATATGAGAGGACGAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5100  
Qy 5449 COTGAAGTTCATCAGGTTCACACCAAGCCTGAGCGCCAGTGGACACCAACCAATGT 5508  
Db 5101 COTGAAGTTCATCAGGTTCACACCAAGCCTGAGCGCCAGTGGACACCAACCAATGT 5160  
Qy 5509 TCAGCTCTACTGATATCCAGTGCAGTGCACCCCAAGGAGAGGAGGAGGAGGAGGAGGAGG 5568  
Db 5161 TCAGCTCTACTGATATCCAGTGCAGTGCACCCCAAGGAGAGGAGGAGGAGGAGGAGGAGG 5220  
Qy 5569 AATCAACCTTGTCTCTGACAGTCACTCGTGGTGTGATCAGGACTTATGTTGGGCCACAA 5628  
Db 5221 AATCAACCTTGTCTCTGACAGTCACTCGTGGTGTGATCAGGACTTATGTTGGGCCACAA 5280

QY 5629 ATATGAAGTGAAGTCTATCTCTTAAAGACACTTTTGAACAAGACAGACAGCTCAGGAGT 5688  
DB ATATGAAGTGAAGTCTATCTCTTAAAGACACTTTTGAACAAGACAGACAGCTCAGGAGT 5340  
QY 5689 TGTCAACCACTCTGGGAATGTCTAGCCCAAGACAGAGAGGCTCTGTGACAGATGCTACTGA 5748  
DB TGTCAACCACTCTGGGAATGTCTAGCCCAAGACAGAGAGGCTCTGTGACAGATGCTACTGA 5400  
QY 5749 GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGA 5808  
DB GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGA 5460  
QY 5809 TGCCTGTGAGCAATGGCCAGACTCCATTCAGAGAACCATCAAGCCAGATGTCAGAA 5868  
DB TGCCTGTGAGCAATGGCCAGACTCCATTCAGAGAACCATCAAGCCAGATGTCAGAA 5520  
QY 5869 CTACACCATCCTGGCTTAAACACCAAGCACTGACTACAGAGTCTACCTGTGACACCTTGA 5928  
DB CTACACCATCCTGGCTTAAACACCAAGCACTGACTACAGAGTCTACCTGTGACACCTTGA 5580  
QY 5929 TGACAAATGCTCGGAGCTCCCTGTGTGTCTGAGCGCTTCACTGGCTTCAAGCCATTCGACCAATC 5988  
DB TGACAAATGCTCGGAGCTCCCTGTGTGTCTGAGCGCTTCACTGGCTTCAAGCCATTCGACCAATC 5640  
QY 5989 CAACCTGCGTCTTCTGGCCACCAACCCCAATTCCTTGTGTGTATCATGCGAGCCGCCACG 6048  
DB CAACCTGCGTCTTCTGGCCACCAACCCCAATTCCTTGTGTGTATCATGCGAGCCGCCACG 5700  
QY 6049 TGCAGGATTAACGGCTATCATATCAATGATGAGAGGCTTGGCTTCTCTCCAGAGAAAT 6108  
DB TGCAGGATTAACGGCTATCATATCAATGATGAGAGGCTTGGCTTCTCTCCAGAGAAAT 5760  
QY 6109 GGTCCCTCGGCCCGCTGGTGTACAGAGGCTACTATTCTGCGCTTGGAGACCGGGAAC 6168  
DB GGTCCCTCGGCCCGCTGGTGTACAGAGGCTACTATTCTGCGCTTGGAGACCGGGAAC 5820  
QY 6169 CGAATATCAATTTATGTCATTCCTGAGAGAAATATCAGAGAGCGAGCCCTGATTTGG 6228  
DB CGAATATCAATTTATGTCATTCCTGAGAGAAATATCAGAGAGCGAGCCCTGATTTGG 5880  
QY 6229 AAGGAAAGACAGAGAGCTTCCCACTGTAACCTTCCACACCCCAATTCATTCAG 6288  
DB AAGGAAAGACAGAGAGCTTCCCACTGTAACCTTCCACACCCCAATTCATTCAG 5940  
QY 6289 ACCAGAGATCTTGGATGTTTCCCTTCCACAGTTCAAAGAGCCCTTTCGTCACCCACCTGG 6348  
DB ACCAGAGATCTTGGATGTTTCCCTTCCACAGTTCAAAGAGCCCTTTCGTCACCCACCTGG 6000  
QY 6349 GTATGACACTGGAATGATTCAGCTTCTGCGACTTCTGTCAGCAACCCAGTGTGG 6408  
DB GTATGACACTGGAATGATTCAGCTTCTGCGACTTCTGTCAGCAACCCAGTGTGG 6060  
QY 6409 GCAACAAATGATCTTTGAGGAACATGGTTTATAGGCGGACCAACCGCCCAACCGCCAC 6468  
DB GCAACAAATGATCTTTGAGGAACATGGTTTATAGGCGGACCAACCGCCCAACCGCCAC 6120  
QY 6469 CCCATAAGGCTAGGCAAGACCATACCGCGGAATGTAGGAGAGAGCTCTCTCTCA 6528  
DB CCCATAAGGCTAGGCAAGACCATACCGCGGAATGTAGGAGAGAGCTCTCTCTCA 6180  
QY 6529 GACAAACATCTCATGGGCCCACTTCAGAGACACTTCTGAGTACATCTTTCATGTCATCC 6588  
DB GACAAACATCTCATGGGCCCACTTCAGAGACACTTCTGAGTACATCTTTCATGTCATCC 6240  
QY 6589 TGTGGCACTGATGAAGAACCCCTTACGTTACGGGTTCTTGAAATTTCTACAGTGGCCAC 6648  
DB TGTGGCACTGATGAAGAACCCCTTACGTTACGGGTTCTTGAAATTTCTACAGTGGCCAC 6300  
QY 6649 TCTGACAGGCTTCCAGAGGTCACACCTTACACATCATAGTGGAGGCACTGAAAGACCA 6708  
DB TCTGACAGGCTTCCAGAGGTCACACCTTACACATCATAGTGGAGGCACTGAAAGACCA 6360

QY 6709 GCAGAGGCATAAAGTTCGGGAAGAGTGTCTACCGTGGGCAACTCTGTCAACGAAGGCTT 6768  
DB GCAGAGGCATAAAGTTCGGGAAGAGTGTCTACCGTGGGCAACTCTGTCAACGAAGGCTT 6420  
QY 6769 GAACCAACCTTACCGATGACTCGTCTTGAACCCCTACAGAGTCTCCATTTATCCCGTTG 6828  
DB GAACCAACCTTACCGATGACTCGTCTTGAACCCCTACAGAGTCTCCATTTATCCCGTTG 6480  
QY 6829 AGATGAGTGGGAACGAATGTCTCAATCAGGCTTTAACTGTGTGTCAGTGTCTTACGCTT 6888  
DB AGATGAGTGGGAACGAATGTCTCAATCAGGCTTTAACTGTGTGTCAGTGTCTTACGCTT 6540  
QY 6889 TGGAGTGTGTCATTTTCAGATGTGATTCATCTAGATGTGTCATGAGGCAACGTTGTGACCTA 6948  
DB TGGAGTGTGTCATTTTCAGATGTGATTCATCTAGATGTGTCATGAGGCAACGTTGTGACCTA 6600  
QY 6949 CAAGATTGGAGAGAGTGGGACCGCTCAGGAGAGAAATGSCCAGATGATGAGCTGACATG 7008  
DB CAAGATTGGAGAGAGTGGGACCGCTCAGGAGAGAAATGSCCAGATGATGAGCTGACATG 6660  
QY 7009 TCTTGGGAACGGAAGAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTACATGA 7068  
DB TCTTGGGAACGGAAGAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTACATGA 6720  
QY 7069 TGGGAAGACATACCACTAGGAGAACAGTGGGAGAGAAATATCTCGTGCCATTTGCTC 7128  
DB TGGGAAGACATACCACTAGGAGAACAGTGGGAGAGAAATATCTCGTGCCATTTGCTC 6780  
QY 7129 CTGCAATGCTTTGGAGCCAGCGGGCTGGCGCTGTGACAACTGCGGAGACCTGGGG 7188  
DB CTGCAATGCTTTGGAGCCAGCGGGCTGGCGCTGTGACAACTGCGGAGACCTGGGG 6840  
QY 7189 TGAACCCAGTCCCGAAGGCACTACTGCGCAGTCTTACAAACCACTATTCTCAGAGATACCA 7248  
DB TGAACCCAGTCCCGAAGGCACTACTGCGCAGTCTTACAAACCACTATTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAACAAAACACTAATGTTAATTGCCAATGAGTGTCTCATGCCCTTTAGATGTACA 7308  
DB TCAGAGAACAAAACACTAATGTTAATTGCCAATGAGTGTCTCATGCCCTTTAGATGTACA 6960  
QY 7309 GGCTGACAGAGAGAGATCCCGAGAGTAAATCATCTTTCCAACTCCAGAGGAAACAGCATGT 7368  
DB GGCTGACAGAGAGAGATCCCGAGAGTAAATCATCTTTCCAACTCCAGAGGAAACAGCATGT 7020  
QY 7369 CTCTCTGCCAAGATCCATCTTAACTGGAGTGTAGTTCAGAGACCCAGCTTAGAGTTCTTC 7428  
DB CTCTCTGCCAAGATCCATCTTAACTGGAGTGTAGTTCAGAGACCCAGCTTAGAGTTCTTC 7080  
QY 7429 TTTCTTTCTTAAAGCCCTTGTCTGAGGAGTCTCCAGCTTCACTCACTCACTCAGCT 7488  
DB TTTCTTTCTTAAAGCCCTTGTCTGAGGAGTCTCCAGCTTCACTCACTCACTCAGCT 7140  
QY 7489 TCTCCAAGCATACCCCTGGAGTTTCTGAGGTTTCTCATAAATGAGGCTGCACTT 7548  
DB TCTCCAAGCATACCCCTGGAGTTTCTGAGGTTTCTCATAAATGAGGCTGCACTT 7200  
QY 7549 GCCTGTTCTGTTGAGATTTCAATACCGCTCAGTATTTTAAATGAGTGTCTAAGA 7608  
DB GCCTGTTCTGTTGAGATTTCAATACCGCTCAGTATTTTAAATGAGTGTCTAAGA 7260  
QY 7609 TTTGGTTTGGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAAAATGTTTGAAT 7668  
DB TTTGGTTTGGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAAAATGTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAAGTACCCAAACACTTCTGCTTTCACCTAAGTGT 7728  
DB GATATGACCAAAATTTTAAAGTAGGAAAGTACCCAAACACTTCTGCTTTCACCTAAGTGT 7380  
QY 7729 CTGCGCCGCAATCTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCA 7788  
DB CTGCGCCGCAATCTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCA 7440  
QY 7789 CAGTACT 7795









2y 3889 AATTACACAAACCCCTACAAACGCGCAGCAGGAAATCTCTTGGAAAGAGTGCTGCATCC 3948  
2b 3541 AATTACCAACACCCCTACAAACGCGCCAGCAGGAAATCTCTTGGAAAGAGTGCTGCATCC 3600  
2y 3949 TGATCAGAGTCTCTGCACTTTTGAATACCTGAGTCCCGGCTCGAGTACAAATGTCAGTGT 4008  
2b 3601 TGATCAGAGTCTCTGCACTTTTGAATACCTGAGTCCCGGCTCGAGTACAAATGTCAGTGT 3660  
2y 4009 TTACACGTCTCAAGATGACAAAGGAAAGTCCCTATCTCTGTATACCATCATCCAGCTGT 4068  
2b 3661 TTACACGTCTCAAGATGACAAAGGAAAGTCCCTATCTCTGTATACCATCATCCAGCTGT 3720  
2y 4069 TCCTCTCCCACTGACCTGCGATTTACCAACAATTTGTCAGACACCATGCGTGTACCTG 4128  
2b 3721 TCCTCTCCCACTGACCTGCGATTTACCAACAATTTGTCAGACACCATGCGTGTACCTG 3780  
2y 4129 GGCTCCACCCCATCCATTTGATTTACCAACTTCTGCTGCTGCTTACTCACCCTGTGAANA 4188  
2b 3781 GGCTCCACCCCATCCATTTGATTTACCAACTTCTGCTGCTGCTTACTCACCCTGTGAANA 3840  
2y 4189 TGAGGAAGATGTTGTCAGAGTGTCAATTTCTCCTTTCAGACAATGCGAGTGTCTTAACAA 4248  
2b 3841 TGAGGAAGATGTTGTCAGAGTGTCAATTTCTCCTTTCAGACAATGCGAGTGTCTTAACAA 3900  
2y 4249 TCTCTGCTGCTGTCAGAAATATGATGAGTGTCTCAGAGTGTCTCAGAAACAATGAGAG 4308  
2b 3901 TCTCTGCTGCTGTCAGAAATATGATGAGTGTCTCAGAGTGTCTCAGAAACAATGAGAG 3960  
2y 4309 CACACCTCTTTAGAGGAAGACAGAAACAGAGTCTTGTATTTCCCAACTGGCATTTTC 4368  
2b 3961 CACACCTCTTTAGAGGAAGACAGAAACAGAGTCTTGTATTTCCCAACTGGCATTTTC 4020  
2y 4369 TGATATTACTGCGAACTCTTTTACTGTCAGTGTGCTCTCTGAGGCAACCATCTGG 4428  
2b 4021 TGATATTACTGCGAACTCTTTTACTGTCAGTGTGCTCTCTGAGGCAACCATCTGG 4080  
2y 4429 CTACAGATCCGCACTCATCCGAGCACTTCACTGGGAGACCTCGAGAGAGTCCGGTGCC 4488  
2b 4081 CTACAGATCCGCACTCATCCGAGCACTTCACTGGGAGACCTCGAGAGAGTCCGGTGCC 4140  
2y 4489 CCACTCTCGGAATTCATCACTCCACCACTCACTCCAGGCAAGAGTATGCTGAC 4548  
2b 4141 CCACTCTCGGAATTCATCACTCCACCACTCACTCCAGGCAAGAGTATGCTGAC 4200  
2y 4549 CATCGTGTCTTTAATGCGAGAGGAAAGTCCCTATTGATTTGGCCCAACATCAACAGT 4608  
2b 4201 CATCGTGTCTTTAATGCGAGAGGAAAGTCCCTATTGATTTGGCCCAACATCAACAGT 4260  
2y 4609 TTCTGATGTTCCGAGGAGCTGGAAGTTGTTGCTGGAGACCCCAAGCTCTAGTACAG 4668  
2b 4261 TTCTGATGTTCCGAGGAGCTGGAAGTTGTTGCTGGAGACCCCAAGCTCTAGTACAG 4320  
2y 4669 CTGGAGTCTCTGCTGTCACTGAGTATATACAGATCACTTACGGAGAAACAGAGG 4728  
2b 4321 CTGGAGTCTCTGCTGTCACTGAGTATATACAGATCACTTACGGAGAAACAGAGG 4380  
2y 4729 AAATAGCCCTGTCAGAGTTCATCTGCTGGAGCAAGTCTACAGCTACCATCAGCG 4788  
2b 4381 AAATAGCCCTGTCAGAGTTCATCTGCTGGAGCAAGTCTTACAGCTACCATCAGCG 4440  
2y 4789 CTTTAAACCTGGAGTTGATTTATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGAC 4848  
2b 4441 CTTTAAACCTGGAGTTGATTTATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGAC 4500  
2y 4849 CCCGCAAGCAGCAAGCAATTTCCATTATTTACCGAAACAGAAATTTGACAAACATCCCA 4908  
2b 4501 CCCGCAAGCAGCAAGCAATTTCCATTATTTACCGAAACAGAAATTTGACAAACATCCCA 4560  
2y 4909 GATGCAAGTGAACGATGTTGAGCAACAGCAATGATGTCAGAGTGTCTGCTTCAAGTTTC 4968  
2b 4561 GATGCAAGTGAACGATGTTGAGCAACAGCAATGATGTCAGAGTGTCTGCTTCAAGTTTC 4620

QY 4969 CCCTGTTACTGTTTACAGAGTAAACCACTCCCAAAATGGAACAGCAGCAACCAAAAC 5028  
Db 4621 CCCTGTTACTGTTTACAGAGTAAACCACTCCCAAAATGGAACAGCAGCAACCAAAAC 4680  
QY 5029 TAAACTGCAAGTCCAGATCAAAACAGAAATGATTTGAAGGCTTGACGCCACAGTGA 5088  
Db 4681 TAAACTGCAAGTCCAGATCAAAACAGAAATGATTTGAAGGCTTGACGCCACAGTGA 4740  
QY 5089 GTATGTTGTTAGTGTCTATGCTCAGAAATCCAAAGCGGAGAGTGTGAGTCTGTTTCAGAC 5148  
Db 4741 GTATGTTGTTAGTGTCTATGCTCAGAAATCCAAAGCGGAGAGTGTGAGTCTGTTTCAGAC 4800  
QY 5149 TGCAGTAAACAAATGATCGCCCTTAAAGACTGGCATTTCACTGATGTGATGTCGATTC 5208  
Db 4801 TGCAGTAAACAAATGATCGCCCTTAAAGACTGGCATTTCACTGATGTGATGTCGATTC 4860  
QY 5209 CATCAAAATTCCTTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 5268  
Db 4861 CATCAAAATTCCTTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 4920  
QY 5269 GAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACTGATGTGATGGAAGACACTGC 5328  
Db 4921 GAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACTGATGTGATGGAAGACACTGC 4980  
QY 5329 AGAGTGCAGAGGCTCAGACCGGTTCTGAGTACACAGTCACTGCTGCTTGCCTTGCACGA 5388  
Db 4981 AGAGTGCAGAGGCTCAGACCGGTTCTGAGTACACAGTCACTGCTGCTTGCCTTGCACGA 5040  
QY 5389 TGATATGAGAGGAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTCTGCACCAACTGA 5448  
Db 5041 TGATATGAGAGGAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTCTGCACCAACTGA 5100  
QY 5449 CCTGAGTCTACTCAGTCAACACCAAGCTCAGCGCCAGTGGACACCAACCAATGT 5508  
Db 5101 CCTGAGTCTACTCAGTCAACACCAAGCTCAGCGCCAGTGGACACCAACCAATGT 5160  
QY 5509 TCAGTCTACTGATATCGAGTCCGGTTCACCCCAAGAGAACCGGACCAATGAAGA 5568  
Db 5161 TCAGTCTACTGATATCGAGTCCGGTTCACCCCAAGAGAACCGGACCAATGAAGA 5220  
QY 5569 AATCAACCTTCTCTGACAGCTCATCCGTGGTGTATTCAGGACTTATGTTGGCCACCAA 5628  
Db 5221 AATCAACCTTCTCTGACAGCTCATCCGTGGTGTATTCAGGACTTATGTTGGCCACCAA 5280  
QY 5629 ATATGAAGTGTGTCTATGCTCTTAAAGGACACTTTGCAAGCAGACAGCTCAGGAGT 5688  
Db 5281 ATATGAAGTGTGTCTATGCTCTTAAAGGACACTTTGCAAGCAGACAGCTCAGGAGT 5340  
QY 5689 TGTCACTCTGAGAGTGTGAGCCCAAGAGGGCTGCTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCACTCTGAGAGTGTGAGCCCAAGAGGGCTGCTGTGACAGATGCTACTGA 5400  
QY 5749 GACCACTATCATTAGCTGGAGAACCAAGACTGAGACGATCACTGCTCCAGTTGA 5808  
Db 5401 GACCACTATCATTAGCTGGAGAACCAAGACTGAGACGATCACTGCTCCAGTTGA 5460  
QY 5809 TGCCGTTCCAGCCCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAAG 5868  
Db 5461 TGCCGTTCCAGCCCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAAG 5520  
QY 5869 CTACACCATCTAGCTGCTTAAACAGGCACTGACTCAAGATCTACCTGTACCTTTGAA 5928  
Db 5521 CTACACCATCTAGCTGCTTAAACAGGCACTGACTCAAGATCTACCTGTACCTTTGAA 5580  
QY 5929 TGCAATCTCGAGCTCCCTGCTGCTCATCGAGCCCTCCACTGCCATTTGATGACCATC 5988  
Db 5581 TGCAATCTCGAGCTCCCTGCTGCTCATCGAGCCCTCCACTGCCATTTGATGACCATC 5640  
QY 5989 CAACTGCTGCTTCTGGCCACCAACCAATTTCTGCTGTATCATGAGCGCCGACG 6048  
Db 5641 CAACTGCTGCTTCTGGCCACCAACCAATTTCTGCTGTATCATGAGCGCCGACG 5700  
QY 6049 TGCCAGGATTACCGGCTACATCATCAAGATGAGAAAGCTCGGGTCTCTCTCCAGAGAGT 6108



ATTORNEY/AGENT INFORMATION:

NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: GOLD=1A PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

CT-US95-09819-6

Query Match 95.2%; Score 7423; DB 5; Length 7680;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

349	GAAGACAGAGCAGGCTCAGCAATGGTTGACCCAGTCCCGGTGGCTGTGCTGCA	408
1	GAAGACAGAGCAGGCTCAGCAATGGTTGACCCAGTCCCGGTGGCTGTGCTGCA	60
409	AAGCAAGCCCGTTGTTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGAGG	468
61	AGCAAGCCCGTTGTTATGACAAATGGAACCACTATCAGATAAATCAACAGTGGAGG	120
469	GACCTACCTAGGCAATGCGTTGGTTGTTGTTATGAGGAAAGCCGAGGTTTAACTG	528
121	GACCTACCTAGGCAATGCGTTGGTTGTTGTTATGAGGAAAGCCGAGGTTTAACTG	180
529	CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTGTAAGAGTACACTGGGAACACTTACG	588
181	CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTGTAAGAGTACACTGGGAACACTTACG	240
589	AGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGTCATCG	648
241	AGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGTCATCG	300
649	GGCTGGCGGAGGAGATGAAGCTGTACCATCCCAACCGCTGCCATGAAGGGGTCAGTC	708
301	GGCTGGCGGAGGAGATGAAGCTGTACCATCCCAACCGCTGCCATGAAGGGGTCAGTC	360
709	CTACAAGATTGGTGACACCTTGGAGGAGACCATGAGACTGGTGTGTTACATGTTAGAGTG	768
361	CTACAAGATTGGTGACACCTTGGAGGAGACCATGAGACTGGTGTGTTACATGTTAGAGTG	420
769	TGTGTCTTGTGTAATGGAAGAGAGATGGACCTGCAAGCCCATAGCTGAGAAGTGT	828
421	TGTGTCTTGTGTAATGGAAGAGAGATGGACCTGCAAGCCCATAGCTGAGAAGTGT	480
829	TGATCATGCTGCTGGGACTTCTTATGCTGCGAGAGAAAGCTGGAGAAAGCCCTACCAAG	888
481	TGATCATGCTGCTGGGACTTCTTATGCTGCGAGAGAAAGCTGGAGAAAGCCCTACCAAG	540
889	CTGGATGATGTTAGATTGACTTGGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	948
541	CTGGATGATGTTAGATTGACTTGGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600
949	TAGAATAGATGCAACGATCAGGACCAAGGACATCTATAGAATTGGAGACACTTGGAG	1008
601	TAGAATAGATGCAACGATCAGGACCAAGGACATCTATAGAATTGGAGACACTTGGAG	660
1009	CAAGAAGGATAATCGAGGAACTCTGCTCCAGTGCATCTCCAGGCAACCGCGGAGGAGA	1068
661	CAAGAAGGATAATCGAGGAACTCTGCTCCAGTGCATCTCCAGGCAACCGCGGAGGAGA	720
1069	GTGGAAGTGTGAGGAGCAGACCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1128
721	GTGGAAGTGTGAGGAGCAGACCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	780

QY	1129	CGATGTTCTGTGAGCTGTTTACCAACCGCAGCTCACCCTCAGCTCCTCCTATGAGCCA	1188
Db	781	CGATGTTCTGTGAGCTGTTTACCAACCGCAGCTCACCCTCAGCTCCTCCTATGAGCCA	840
QY	1189	CTGTGTACAGACAGTGTGTGTCTACTCTGTGGGATGAGTGGTGGTGAAGACACAGG	1248
Db	841	CTGTGTACAGACAGTGTGTGTCTACTCTGTGGGATGAGTGGTGGTGAAGACACAGG	900
QY	1249	AAATAAGCAAAATGCTTTGCAAGCTGCTGGCAACGAGTCAAGTGGCAAGACAGCTGT	1308
Db	901	AAATAAGCAAAATGCTTTGCAAGCTGCTGGCAACGAGTCAAGTGGCAAGACAGCTGT	960
QY	1309	AACCCAGACTTACGTTGCAAGTCAAAATGAGAGGCAATGCTTACATTCACCTACAA	1368
Db	961	AACCCAGACTTACGTTGCAAGTCAAAATGAGAGGCAATGCTTACATTCACCTACAA	1020
QY	1369	TGGCAGGACGTTCTACTCTGCAACCAAGAGGCGCAGAGGAGGACATCTTTGGTGCAG	1428
Db	1021	TGGCAGGACGTTCTACTCTGCAACCAAGAGGCGCAGAGGAGGACATCTTTGGTGCAG	1080
QY	1429	CACAACTTCGAATATGAGCAGGACGAGAAATCTCTTTCTGCAAGACCACTGTTTT	1488
Db	1081	CACAACTTCGAATATGAGCAGGACGAGAAATCTCTTTCTGCAAGACCACTGTTTT	1140
QY	1489	GGTTTCAGACTCGAGGAGGAAATTCCAATGCTGCTTGTGCCACTTCCCTTCTATACAA	1548
Db	1141	GGTTTCAGACTCGAGGAGGAAATTCCAATGCTGCTTGTGCCACTTCCCTTCTATACAA	1200
QY	1549	CAACCACTTACACTGATTCGACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1608
Db	1201	CAACCACTTACACTGATTCGACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1260
QY	1609	GACCACACAGAACTATGATCCGACCAAGAAATTTGGGTTCTGCCCCATGCTGCCACGA	1668
Db	1261	GACCACACAGAACTATGATCCGACCAAGAAATTTGGGTTCTGCCCCATGCTGCCACGA	1320
QY	1669	GGAAATCTGCAACCAATGAAGGGTCACTGATCCGATTTGGAGATCAGTGGGATAGCA	1728
Db	1321	GGAAATCTGCAACCAATGAAGGGTCACTGATCCGATTTGGAGATCAGTGGGATAGCA	1380
QY	1729	GCATGACATGGTGCACATGATGAGTGCACGTGTGTTGGAAATGGTCTGGGAGATGGAC	1788
Db	1381	GCATGACATGGTGCACATGATGAGTGCACGTGTGTTGGAAATGGTCTGGGAGATGGAC	1440
QY	1789	ATGCATGCTCTACTCGCAGCTTCAGATCAGTGCATGTTGATGACATCACTTACAAATG	1848
Db	1441	ATGCATGCTCTACTCGCAGCTTCAGATCAGTGCATGTTGATGACATCACTTACAAATG	1500
QY	1849	GAACGACACATTCACCAAGCGTCAATGAAGAGGGGACATGCTGAACTGTACATGCTCGG	1908
Db	1501	GAACGACACATTCACCAAGCGTCAATGAAGAGGGGACATGCTGAACTGTACATGCTCGG	1560
QY	1909	TCAGGGTGGGCGCAGGTGGAAGTGTGATCCCGTGCAGCAATGCCAGATTCAGAGACTGG	1968
Db	1561	TCAGGGTGGGCGCAGGTGGAAGTGTGATCCCGTGCAGCAATGCCAGATTCAGAGACTGG	1620
QY	1969	GACGTTTTATCAAAATGGAATTCATGGGAGAGATGTCATGGTGTACATGATGACAGTG	2028
Db	1621	GACGTTTTATCAAAATGGAATTCATGGGAGAGATGTCATGGTGTACATGATGACAGTG	1680
QY	2029	CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAACCTTTTACAGACCTTATCAAAG	2088
Db	1681	CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAACCTTTTACAGACCTTATCAAAG	1740
QY	2089	CTCAAGTGGTCTGTGCAAGTATTTATCACTGAGATCCCGAGTCCAGCCCACTCCACCC	2148
Db	1741	CTCAAGTGGTCTGTGCAAGTATTTATCACTGAGATCCCGAGTCCAGCCCACTCCACCC	1800
QY	2149	CATCCAGTGAATGCACACAGCCATCTCACTTCCAAAGTACATCTCAGGTGGAGACC	2208
Db	1801	CATCCAGTGAATGCACACAGCCATCTCACTTCCAAAGTACATCTCAGGTGGAGACC	1860
QY	2209	TAAAAATTTCTGTAGGCGCTTGGAAAGGAGGCTACCATACAGGCGGCACTTAAACTCTTACAC	2268

1861 TAAAAATTTCTAGGCGGTGGAGAAAGCTACCATACAGGCCACTTTAAACTCCTACAC 1920  
2269 CATCAAGGCTGAAGCCTGGTGTGGTATACAGGGCCAGCTCATCAGCATCCAGCAGTA 2328  
1921 CATCAAGGCTGAAGCCTGGTGTGGTATACAGGGCCAGCTCATCAGCATCCAGCAGTA 1980  
2329 CGGCCACCAAGAGTGAATCGGTTTGACTTCCACACACACACACACACACACACACAC 2308  
1981 CGGCCACCAAGAGTGAATCGGTTTGACTTCCACACACACACACACACACACACACAC 2040  
2389 CAGCAACACCGGTGACAGGAGAGACGACCTCCCTTTTCTCTCTCTCTCTCTCTCTCT 2448  
2041 CAGCAACACCGGTGACAGGAGAGACGACCTCCCTTTTCTCTCTCTCTCTCTCTCTCT 2100  
2449 TGTGACCGAAATCAAGCCAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 2508  
2101 TGTGACCGAAATCAAGCCAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 2160  
2509 GTCCGGATTCCTGGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2568  
2161 GTCCGGATTCCTGGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2220  
2569 TCTTCCAGCAGACGACCTTCTGTGAAATCCCTGACCTGCTTCTGTGCTTCTGTCTTCA 2628  
2221 TCTTCCAGCAGACGACCTTCTGTGAAATCCCTGACCTGCTTCTGTGCTTCTGTCTTCA 2280  
2629 TGTAAATGTCTATCAGATATCTGAGATGGGAGGAGAGTGTGATCTCTTCTTCTTCA 2688  
2281 TGTAAATGTCTATCAGATATCTGAGATGGGAGGAGAGTGTGATCTCTTCTTCTTCA 2340  
2689 AACACAGCGCTGATGTGCTTCTGACCGAGCTGTGACCAAGTGTGATGACCTCAAT 2748  
2341 AACACAGCGCTGATGTGCTTCTGACCGAGCTGTGACCAAGTGTGATGACCTCAAT 2400  
2749 TCTTGTGTGCTGGAGCAGACCCAGGCTCCCATCACAGGCTACAGATAGTCTATTCGCC 2808  
2401 TCTTGTGTGCTGGAGCAGACCCAGGCTCCCATCACAGGCTACAGATAGTCTATTCGCC 2460  
2809 ATCAGTAGAGTGTAGCAGACAGAACTCAACTTCTCTGAAACTGCAAACTCCGTCACCT 2868  
2461 ATCAGTAGAGTGTAGCAGACAGAACTCAACTTCTCTGAAACTGCAAACTCCGTCACCT 2520  
2869 CAGTGACTTGAACCTGTGTGTGATATACATCACTATCTATCTGTGTGGAAGAAATCA 2928  
2521 CAGTGACTTGAACCTGTGTGTGATATACATCACTATCTATCTGTGTGGAAGAAATCA 2580  
2929 AGAAAGTACACCTGTGTGATATCAACAGAAACCACTGGCACCACCGCTCAGATACAGT 2988  
2581 AGAAAGTACACCTGTGTGATATCAACAGAAACCACTGGCACCACCGCTCAGATACAGT 2640  
2989 GCCCTCTCCAGGACCTGCAGTTTGTGGAAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 3048  
2641 GCCCTCTCCAGGACCTGCAGTTTGTGGAAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 2700  
3049 GACACCGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3108  
2701 GACACCGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2760  
3109 TGGCAGCAGCGGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3168  
2761 TGGCAGCAGCGGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2820  
3169 GTCCCTCTGGGTGACCTATTACTTCAAAAGTCTTTTGTGAGTGTGAGTGTGAGTGTGAG 3228  
2821 GTCCCTCTGGGTGACCTATTACTTCAAAAGTCTTTTGTGAGTGTGAGTGTGAGTGTGAG 2880  
3229 GGCTCTGACTGTCAACAGACAAACAACTGGAGTGTGCTTCTTCTTCTTCTTCTTCTTCT 3288  
2881 GGCTCTGACTGTCAACAGACAAACAACTGGAGTGTGCTTCTTCTTCTTCTTCTTCTTCT 2940  
3289 TGAAGTGTGACTGTCTGTGCTGAGATGGTCTCCACTCGGGCCAGATTAACAGGATA 3348

2941 TGAACCTGATTTCTATGTCTCTGTGAGATGGACTCCACTCGGGCCAGATTAACAGGATA 3000  
3349 CCGACTGACCGTGGGCTTTACCCGAAGAGGAGCAGCCAGCAGTACAAATGTGGGTCCCTC 3408  
3001 CCGACTGACCGTGGGCTTTACCCGAAGAGGAGCAGCCAGCAGTACAAATGTGGGTCCCTC 3060  
3409 TGTCTCCAAAGTACCCACTGAGGAATCTCAGCCTGCATCTGAGTACACCGTATCCCTGT 3468  
3061 TGTCTCCAAAGTACCCCTCTGAGGAATCTGAGCCTGCATCTGAGTACACCGTATCCCTGT 3120  
3469 GGCATATAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTTACCACTCAGCC 3528  
3121 GGCATATAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTTACCACTCAGCC 3180  
3529 TGGGAGCTCTANTTCCACTTCAACACACGAGGTGACTGAGACCACTATGTGATCACATG 3588  
3181 TGGGAGCTCTANTTCCACTTCAACACACGAGGTGACTGAGACCACTATGTGATCACATG 3240  
3589 GAGCCTGTCTCAAGAAATTTGGTTTAAAGCTGGGTGTACGACCAAGCCAGGGAGGAGGC 3648  
3241 GAGCCTGTCTCAAGAAATTTGGTTTAAAGCTGGGTGTACGACCAAGCCAGGGAGGAGGC 3300  
3649 ACCACAGAGAGTGTGACTTCACTCAGACCTCAGGAAGCATGTTGTTCGGCTTGACTCCAGGAGT 3708  
3301 ACCACAGAGAGTGTGACTTCACTCAGGAAGCATGCTGTGTCCGGCTTGACTCCAGGAGT 3360  
3709 AGAATACGTCTACACCATCCAAAGTCTCTGAGAGATGGACAGGAAGAGATGCGCCAAATGT 3768  
3361 AGAATACGTCTACACCATCCAAAGTCTCTGAGAGATGGACAGGAAGAGATGCGCCAAATGT 3420  
3769 AAACAAAGTGTGACACCAATTTGTCTCCACCAACAACTTGGATCTGGAGGCAAAACCTTGA 3828  
3421 AAACAAAGTGTGACACCAATTTGTCTCCACCAACAACTTGGATCTGGAGGCAAAACCTTGA 3480  
3829 CACTGAGTGTCTACAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTTGGTTATAG 3888  
3481 CACTGAGTGTCTACAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTTGGTTATAG 3540  
3889 AATTACCAACACCCCTTACAAACCGGCGAGGAGAAATTTCTTGGAGAAAGTGGTCCATGC 3948  
3541 AATTACCAACACCCCTTACAAACCGGCGAGGAGAAATTTCTTGGAGAAAGTGGTCCATGC 3600  
3949 TGATCAGAGCTCTGACCTTTTGTAAACCTGAGTCCGSCCTGGAGTACAAATGTCACTGT 4008  
3601 TGATCAGAGCTCTGACCTTTTGTAAACCTGAGTCCGSCCTGGAGTACAAATGTCACTGT 3660  
4009 TTACTCTGCAAGGATGACAAAGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
3661 TTACTCTGCAAGGATGACAAAGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
4069 TCTCTCTCCACTGACCTGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTGT 4128  
3721 TCTCTCTCCACTGACCTGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTGT 3780  
4129 GGTCTCCACCCCATCCATTTTAAACCAACTTCTGGTGGCTTACTCACCTGTGAAAAA 4188  
3781 GGTCTCCACCCCATCCATTTTAAACCAACTTCTGGTGGCTTACTCACCTGTGAAAAA 3840  
4189 TGAAGAGATGTGAGAGTGTCTCAATTTCTCTTTCAGACATGAGTGTCTTAAACAA 4248  
3841 TGAAGAGATGTGTGAGAGTGTCTCAATTTCTCTTTCAGACATGAGTGTCTTAAACAA 3900  
4249 TCTCTCTGCTGTGTACAGATATGTAGTGTGTCTCCAGTGTCTACGAACCAACATGAGAG 4308  
3901 TCTCTCTGCTGTGTACAGATATGTAGTGTGTCTCCAGTGTCTACGAACCAACATGAGAG 3960  
4309 CACACCTCTTAGAGGAGACAGAAACAGGTCTTGAATTTCCCAACTGGCATTTGCTTTC 4368  
3961 CACACCTCTTAGAGGAGACAGAAACAGGTCTTGAATTTCCCAACTGGCATTTGCTTTC 4020  
4369 TGAATATTCTGCGCACTCTTTTACTGTGTGATTTGCTTCTCGAGGCCACCATCACTGG 4428  
4021 TGAATATTCTGCGCACTCTTTTACTGTGTGATTTGCTTCTCGAGGCCACCATCACTGG 4080





Db 6241 TGTGGCACTGATGAAGAACCTTACAGTTGAGGTTCTCGAACTTCTACCACTGTCAC 6300  
QY 6649 TCTGACAGCCTTCAACAGAGTGCACCTCAACAATCATAGTGAAGGACCTGAAAGACCA 6708  
Db 6301 TCTGACAGCCTTCAACAGAGTGCACCTCAACAATCATAGTGAAGGACCTGAAAGACCA 6360  
QY 6709 GCAGAGGATAGGTTGCGGAGAGGTTGTTACCGTGGGCAACTCTGTCACAGAGGCTT 6768  
Db 6361 GCAGAGGATAGGTTGCGGAGAGGTTGTTACCGTGGGCAACTCTGTCACAGAGGCTT 6420  
QY 6769 GAACCAACTACGGATGACTCGTGCTTTGACCCCTPACACAGTGTCCCATATATGCGTTGG 6828  
Db 6421 GAACCAACTACGGATGACTCGTGCTTTGACCCCTPACACAGTGTCCCATATATGCGTTGG 6480  
QY 6829 AGATGAGTGGGAACCAATGCTGTAATCAGGCTTTAACTGTTGTCAGTGTCTTAGGCTT 6888  
Db 6481 AGATGAGTGGGAACCAATGCTGTAATCAGGCTTTAACTGTTGTCAGTGTCTTAGGCTT 6540  
QY 6889 TGAAGTGGTCACTTTACAGATGTGATTCATCTAGATGTTGTCATGCAATGTTGTAACCTA 6948  
Db 6541 TGAAGTGGTCACTTTACAGATGTGATTCATCTAGATGTTGTCATGCAATGTTGTAACCTA 6600  
QY 6949 CAAGATTGGAGAGTGGGACCTGACGGAGAAATGCGCAGATGATGAGTGCACATG 7008  
Db 6601 CAAGATTGGAGAGTGGGACCTGACGGAGAAATGCGCAGATGATGAGTGCACATG 6860  
QY 7009 TCTTGGGAACGAAAGAGAGATTCACAGTGTGACCTCATGAGCAACGTTGTACGATGA 7068  
Db 6661 TCTTGGGAACGAAAGAGAGATTCACAGTGTGACCTCATGAGCAACGTTGTACGATGA 6720  
QY 7069 TGGGAGACATACACCTAGGAGAACAGTGCAGAGAGAAATCTCGTGCCATTGCTC 7128  
Db 6721 TGGGAGACATACACCTAGGAGAACAGTGCAGAGAGAAATCTCGTGCCATTGCTC 6780  
QY 7129 CTGCACATGCTTTGAGGACCGAGGCTGCGGTGTGACAACTGCGCAGACCTGCGGG 7188  
Db 6781 CTGCACATGCTTTGAGGACCGAGGCTGCGGTGTGACAACTGCGCAGACCTGCGGG 6940  
QY 7189 TGAACCCAGTCCGAGAGCACTACTGCGGAGTCTCAACAGATTTCTCAGAGATACCA 7248  
Db 6841 TGAACCCAGTCCGAGAGCACTACTGCGGAGTCTCAACAGATTTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTCATGCTTTAGATGTACA 7308  
Db 6901 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTCATGCTTTAGATGTACA 6960  
QY 7309 GGCTGACAGAGAGATTCGAGAGTAAATCATCTTTCCAAATCCAGAGAACAAAGCATGT 7368  
Db 6961 GGCTGACAGAGAGATTCGAGAGTAAATCATCTTTCCAAATCCAGAGAACAAAGCATGT 7020  
QY 7369 CTCTCTGCAAGATCCATCTAAACTGAGTGTATGACAGACCCAGCTTAGAGTTCTTC 7428  
Db 7021 CTCTCTGCAAGATCCATCTAAACTGAGTGTATGACAGACCCAGCTTAGAGTTCTTC 7080  
QY 7429 TTTCTTTCTTAAGCCCTTTGCTCTGAGAGAGTGTCTCCAGCTTCAGCTCAACTCAGAGCT 7488  
Db 7081 TTTCTTTCTTAAGCCCTTTGCTCTGAGAGAGTGTCTCCAGCTTCAGCTCAACTCAGAGCT 7140  
QY 7489 TCTCAAGCATCACCCCTGGAGTGTCTTCCCTGAGGTTTCTCATATAATGAGGCTGCACATT 7548  
Db 7141 TCTCAAGCATCACCCCTGGAGTGTCTTCCCTGAGGTTTCTCATATAATGAGGCTGCACATT 7200  
QY 7549 GCCTGTCTGCTTCAAGTATTCATAACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7608  
Db 7201 GCCTGTCTGCTTCAAGTATTCATAACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7260  
QY 7609 TTTGTTTGGGATCAATAGGAGAGATATGAGGCAACCAAGATGCAAAATGTTTTGAAT 7668  
Db 7261 TTTGTTTGGGATCAATAGGAGAGATATGAGGCAACCAAGATGCAAAATGTTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACATCTTGTCTTTTCACTTAAGTGT 7728

Db 7321 GATATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACACTTCTGCTTTCACTTAAGTGT 7380  
QY 7729 CTGGCCCGCAATACTGTTAGGAAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7788  
Db 7381 CTGGCCCGCAATACTGTTAGGAAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7440  
QY 7789 CAGTACT 7795  
Db 7441 CAGTACT 7447  
RESULT 5  
US-08-259-569-16  
; Sequence 16, Application US/08259569  
; Patent No. 5679320  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Tikva  
; APPLICANT: Levanon, Avigdor  
; APPLICANT: Werber, Moshe  
; APPLICANT: Guy, Rachel  
; APPLICANT: Panet, Amos  
; APPLICANT: Hartman, Jacob  
; APPLICANT: Shaked, Hadassa  
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND  
; TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/259,569  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 703,842  
; FILING DATE: 21-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-977-9550  
; TELEFAX: 212-664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7705 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 5..7681  
; US-08-259-569-16  
Query Match 95.2%; Score 7418.2; DB 1; Length 7705;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7429; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
QY 349 GAAGAGCAAGAGGCGAGGTCTAGCAAAATGTTTCAGCCCGTCCCGTGGCTGTCAATCA 408





2569	QY	TCTTCCAGGCA	GAGCCAC	TTCTGTGTAAACAT	CCCTGACCT	CTGCTCTGCTGGCGGAAATACAT	2628
2221	DB	TCTTCCAGGCA	GAGCCAC	TTCTGTGTAAACAT	CCCTGACCT	CTGCTCTGCTGGCGGAAATACAT	2280
2629	QY	TGTAAATGTCTAT	CAGATATCT	GAGGATCGGGAG	CAGAGTTTGAT	CGCTCTCTACTTCACA	2688
2281	DB	TGTAAATGTCTAT	CAGATATCT	GAGGATCGGGAG	CAGAGTTTGAT	CGCTCTCTACTTCACA	2340
2689	QY	AACAAAGCGCCTGAT	GCCTCTGAC	CCGACTGTGGACCA	AGTTGATGAC	CACTCAAT	2748
2341	DB	AACAAAGCGCCTGAT	GCCTCTGAC	CCGACTGTGGACCA	AGTTGATGAC	CACTCAAT	2400
2749	QY	TGTTGTTTCCGT	CGAGCAGAC	CCCCAGGCT	CCCCATCAGAGG	GATCAGATAATAGTCTATTTCGCC	2808
2401	DB	TGTTGTTTCCGT	CGAGCAGAC	CCCCAGGCT	CCCCATCAGAGG	GATCAGATAATAGTCTATTTCGCC	2460
2809	QY	ATCAGTAGAAGGTAG	CAGACAGAACT	CAACCTTCTGTGAA	CTGGAAATCTCGGTCACT	2868	
2461	DB	ATCAGTAGAAGGTAG	CAGACAGAACT	CAACCTTCTGTGAA	CTGGAAATCTCGGTCACT	2520	
2869	QY	CAGTGACTTTGCA	ACCTGGTTC	CAGTATTAACATCACT	ATCTATGTGTGGAA	AGAAATCA	2928
2521	DB	CAGTGACTTTGCA	ACCTGGTTC	CAGTATTAACATCACT	ATCTATGTGTGGAA	AGAAATCA	2580
2929	QY	AGAAAGTACACCTG	TTGTCACTT	CAACAGAAACCACT	GGCACCCCACTGATACAT	2988	
2581	DB	AGAAAGTACACCTG	TTGTCACTT	CAACAGAAACCACT	GGCACCCCACTGATACAT	2640	
2989	QY	GCCTCTCCAGGGAC	CTGCAGTTT	TGGAAGTGACAG	ACGTGAAGGTCA	CACTCATGTG	2700
2641	DB	GCCTCTCCAGGGAC	CTGCAGTTT	TGGAAGTGACAG	ACGTGAAGGTCA	CACTCATGTG	2600
3049	QY	GACACCGCTGAGAG	TGCAGTGC	CGGTACCGTGTGG	ATGTGATCCCGTCA	ACCTGCC	3108
2701	DB	GACACCGCTGAGAG	TGCAGTGC	CGGTACCGTGTGG	ATGTGATCCCGTCA	ACCTGCC	2760
3109	QY	TGGCGAGCAGGGC	AGGCTGCCAT	CAGCAGGAAC	ACTTTGCAGAA	AGTCA	3168
2761	DB	TGGCGAGCAGGGC	AGGCTGCCAT	CAGCAGGAAC	ACTTTGCAGAA	AGTCA	2820
3169	QY	GTCCCTGGGTCAC	CTATTACTTT	CAAGTCTTTGAG	CTGAGCCTGGGAGG	GAGAGCAA	3228
2821	DB	GTCCCTGGGTCAC	CTATTACTTT	CAAGTCTTTGAG	CTGAGCCTGGGAGG	GAGAGCAA	2880
3229	QY	GCTCTGACTGTCT	CAACAGACAA	CCAAAATGG	ATGCTCCAC	TAACTCAAGTTTGTCAA	3288
2881	DB	GCTCTGACTGTCT	CAACAGACAA	CCAAAATGG	ATGCTCCAC	TAACTCAAGTTTGTCAA	2940
3289	QY	TGAACTGATTTCT	GTCTGTGTAGAT	TGGACTCCAC	CTCGGGCCACAGAT	AACAGATA	3348
2941	DB	TGAACTGATTTCT	GTCTGTGTAGAT	TGGACTCCAC	CTCGGGCCACAGAT	AACAGATA	3000
3349	QY	CCGACTGACCTGGG	CCTTACCCG	AAGAGGAC	CAGCCAGGAGTAC	AAATGTGGTCCCTC	3408
3001	DB	CCGACTGACCTGGG	CCTTACCCG	AAGAGGAC	CAGCCAGGAGTAC	AAATGTGGTCCCTC	3060
3409	QY	TGTCTCCAAAGT	ACCCACTGAG	GAATCTGAC	CGCTGATAC	CCGTATCCCTCGT	3468
3061	DB	TGTCTCCAAAGT	ACCCACTGAG	GAATCTGAC	CGCTGATAC	CCGTATCCCTCGT	3120
3469	QY	GGCCATAAAGGG	CAACAGAGAG	CCCCAAG	CCACTTGA	GTCTTTACCACTG	3528
3121	DB	GGCCATAAAGGG	CAACAGAGAG	CCCCAAG	CCACTTGA	GTCTTTACCACTG	3180
3529	QY	TGGGAGCTCTATT	CCACCTTACAA	CAACCGAGGTGA	CTGAGACCA	CCATTTGTATACATG	3588
3181	DB	TGGGAGCTCTATT	CCACCTTACAA	CAACCGAGGTGA	CTGAGACCA	CCATTTGTATACATG	3240
3589	QY	GACGCTGTCTCAA	GAATTTGGTTT	TAACTGGGTGT	ACGACCAAG	CCAGGAGGAGG	3648
3241	DB	GACGCTGTCTCAA	GAATTTGGTTT	TAACTGGGTGT	ACGACCAAG	CCAGGAGGAGG	3300
3649	QY	ACCACAGAGA	AGTGA	CTTCA	CACTCAGGA	AGCATCGTTGTGTCCGGCTTGACTCCAGGAGT	3708

3301	DB	ACCCAGGAAGTGA	CTTCAGACTCAGGAAGCAATCTGTGTCCGGCTTGACTCCAGGAGT	3360
3709	QY	AGAATACGTCTACACCATCAAGT	CCTCAGAGATGCAAGGAAGAGATGCGCCAAATGTT	3768
3361	DB	AGAATACGTCTACACCATCAAGT	CCTCAGAGATGCAAGGAAGAGATGCGCCAAATGTT	3420
3769	QY	AAACAAAGTGGTGACACCATTTGTC	CCACCAACAACCTTGCATCTGGAGGCAAAACCTCGA	3828
3421	DB	AAACAAAGTGGTGACACCATTTGTC	CCACCAACAACAACCTTGCATCTGGAGGCAAAACCTCGA	3480
3829	QY	CACCTGGAGTGTCTACAGTCTCT	CTGGAGAGGAGCAACCCCCAGACATTTACTGGTTATAG	3888
3481	DB	CACCTGGAGTGTCTACAGTCTCT	CTGGAGAGGAGCAACCCCCAGACATTTACTGGTTATAG	3540
3889	QY	AAATTACCAACACCCCTACAAACGGG	CACAGGGAATTTCTTTGGAAGAGTGGTCCCATCG	3948
3541	DB	AAATTACCAACACCCCTACAAACGGG	CACAGGGAATTTCTTTGGAAGAGTGGTCCCATCG	3600
3949	QY	TGATCAGAGCTCCTGCATTTTGATAA	CCCTGAGTCCCGGCCCTGGAGTACAATGTCACTGT	4008
3601	DB	TGATCAGAGCTCCTGCATTTTGATAA	CCCTGAGTCCCGGCCCTGGAGTACAATGTCACTGT	3660
4009	QY	TTTACACTGTCAAGGATGACAAGGAAG	TGTCCTATCTGATACCATCATCCAGCTGT	4068
3661	DB	TTTACACTGTCAAGGATGACAAGGAAG	TGTCCTATCTGATACCATCATCCAGCTGT	3720
4069	QY	TGCTCCTCCCATGACCTGCGANTCAC	CAACAATGGTCAGACACAATGGCTGTCACTTG	4128
3721	DB	TGCTCCTCCCATGACCTGCGANTCAC	CAACAATGGTCAGACACAATGGCTGTCACTTG	3780
4129	QY	GGCTCCACCCCATCCATTTGATTTAA	CCAACTCTCTGGTGGCTTACTCACCTGTGAAAA	4188
3781	DB	GGCTCCACCCCATCCATTTGATTTAA	CCAACTCTCTGGTGGCTTACTCACCTGTGAAAA	3840
4189	QY	TGAGGAAGATGTTGAGAGTGTGTCAAT	TTTCTCCTTCAGACAATGCAATGGTCTTAACAA	4248
3841	DB	TGAGGAAGATGTTGAGAGTGTGTCAAT	TTTCTCCTTCAGACAATGCAATGGTCTTAACAA	3900
4249	QY	TCTCTGCTGGTACAGAAATATGTAGT	GAGTCTCCAGTGCTACGAACAACATCAGAG	4308
3901	DB	TCTCTGCTGGTACAGAAATATGTAGT	GAGTCTCCAGTGCTACGAACAACATCAGAG	3960
4309	QY	CACACTCTTAGAGGAAGACAGAAAA	CAGTCTTGATTTCCCAACTGGCATTTTC	4368
3961	DB	CACACTCTTAGAGGAAGACAGAAAA	CAGTCTTGATTTCCCAACTGGCATTTTC	4020
4369	QY	TGATATTACTGCAACTCTTTTACTGT	GCACTGGATTGCTCCTCAGACCACCATCACTGG	4428
4021	DB	TGATATTACTGCAACTCTTTTACTGT	GCACTGGATTGCTCCTCAGACCACCATCACTGG	4080
4429	QY	CTACAGGATCCGCATCATCCCGAGCA	TTTCTAGTGGGAGCCTCGAGAGATCGGGTCC	4488
4081	DB	CTACAGGATCCGCATCATCCCGAGCA	TTTCTAGTGGGAGCCTCGAGAGATCGGGTCC	4140
4489	QY	CCACTCTCGGAATTCCTATCACCCCT	CACCAACCTCACTCCAGGCACAGAGTATGTGTCAG	4548
4141	DB	CCACTCTCGGAATTCCTATCACCCCT	CACCAACCTCACTCCAGGCACAGAGTATGTGTCAG	4200
4549	QY	CATCTGTGCTTTAATGGCAGAGGAAG	TCCCTTTATGATGGCCACCAATCAACACT	4608
4201	DB	CATCTGTGCTTTAATGGCAGAGGAAG	TCCCTTTATGATGGCCACCAATCAACACT	4260
4609	QY	TTTCTGATGTTCCGAGGGACCTGGAA	GTGTTGCTCGACCCCCCAACGCTTACTGATCAG	4668
4261	DB	TTTCTGATGTTCCGAGGGACCTGGAA	GTGTTGCTCGACCCCCCAACGCTTACTGATCAG	4320
4669	QY	CTGGGATGCTCCTGCTGTACAGT	CGAGATTTACAGGATCACTTACGGGAACACGAGG	4728
4321	DB	CTGGGATGCTCCTGCTGTACAGT	CGAGATTTACAGGATCACTTACGGGAACACGAGG	4380
4729	QY	AAATAGCCCTGTCCAGGAGTTCATCT	GTGCTGGGAGCAAGTCTACAGTCAACATCAGCGG	4788



6949	QY	CAAGATTGGAGAGAGTGGGACCGTCAGGAGAGAAATGGCCAGATGATGAGCTGCACATG	7008
6601	DB	CAAGATTGGAGAGAGTGGGACCGTCAGGAGAGAAATGGCCAGATGATGAGCTGCACATG	6660
7009	QY	TCATTGGGAACGGAAAAAGGAAATTCAGTGTGACCCCTCATGAGGCAACGTTGTACCATGA	7068
6661	DB	TCATTGGGAACGGAAAAAGGAAATTCAGTGTGACCCCTCATGAGGCAACGTTGTACCATGA	6720
7069	QY	TGGGAGAGACATCCACGTAGGAGAAACAGTGGCAGAGAGGAATATCTCGGTGCCAATTTGCTC	7128
6721	DB	TGGGAGAGACATCCACGTAGGAGAAACAGTGGCAGAGAGGAATATCTCGGTGCCAATTTGCTC	6780
7129	QY	CTGCACATGCTTTTGGAGCCACGGGGGCTGGCGCTGTGACAACTGCCGCAGACCTCGGGG	7188
6781	DB	CTGCACATGCTTTTGGAGCCACGGGGGCTGGCGCTGTGACAACTGCCGCAGACCTCGGGG	6840
7189	QY	TGAACCCAGTCGCCAGAGGCACTACTGGCCAGTCCCTACAAACAGTATCTCAGAGATACCA	7248
6841	DB	TGAACCCAGTCGCCAGAGGCACTACTGGCCAGTCCCTACAAACAGTATCTCAGAGATACCA	6900
7249	QY	TCAGAGAACAAAACACTAATGTTAATTGCCCAATTGAGTGTCTCATGCCTTTAGATGTACA	7308
6901	DB	TCAGAGAACAAAACACTAATGTTAATTGCCCAATTGAGTGTCTCATGCCTTTAGATGTACA	6960
7309	QY	GGCTGACAGAGAAGATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGT	7368
6961	DB	GGCTGACAGAGAAGATTCCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAAGCATGT	7020
7369	QY	CTCTCTGCCAAGATCCATCTTAAACTGGAGTGATGTTAGCAAGACCCAGCTTTPAGAGTTCTTC	7428
7021	DB	CTCTCTGCCAAGATCCATCTTAAACTGGAGTGATGTTAGCAAGACCCAGCTTTPAGAGTTCTTC	7080
7429	QY	TTTCTTTTCTTAAGCCCTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT	7488
7081	DB	TTTCTTTTCTTAAGCCCTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCT	7140
7489	QY	TCTCCAGCATACCCCTGGAGTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7548
7141	DB	TCTCCAGCATACCCCTGGAGTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7200
7549	QY	GCCTGTTTCTGCTTCGAAGTATTCAATACCCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7608
7201	DB	GCCTGTTTCTGCTTCGAAGTATTCAATACCCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7260
7609	QY	TTTGGTTTGGGATCAATAGAAAGCATATGCAGCCCAACCAAGATGCAATGTTTTGAAT	7668
7261	DB	TTTGGTTTGGGATCAATAGAAAGCATATGCAGCCCAACCAAGATGCAATGTTTTGAAT	7320
7669	QY	GATATGACAAAAATTTTAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7728
7321	DB	GATATGACAAAAATTTTAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7380
7729	QY	CTGGCCGCGAATCTGTAGGAAACAGCATGATCTTGTTACTGTGATATTTTAATATCCA	7788
7381	DB	CTGGCCGCGAATCTGTAGGAAACAGCATGATCTTGTTACTGTGATATTTTAATATCCA	7440
7789	QY	CAGTACT 7795	
7441	DB	CAGTACT 7447	

RESULT 6  
US-08-826-885-16  
US-08-826-16, Application US/08826885  
; Patent No. 5869616  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Tikva  
; APPLICANT: Levanon, Avigdor  
; APPLICANT: Werber, Moshe  
; APPLICANT: Guy, Rachel  
; APPLICANT: Panet, Amos  
; APPLICANT: Hartman, Jacob  
; APPLICANT: Shaked, Hadassa

TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND  
 TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME  
 NUMBER OF SEQUENCES: 39  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/826,865  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/703,842  
 FILING DATE: 21-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-977-9550  
 TELEFAX: 212-664-0525  
 TELEX: 422523 COOP UI  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7705 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 5..7681  
 US-08-826-885-16

Query Match	95.2%;	Score 7418.2;	DB 2;	Length 7705;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 7429; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0

349	QY	GAAGAGCAAGAGGCAGGCTCAGCAAAATGGTTACAGCCCAAGTCCC CGGTGGCTGT CAGTCA	408
	Db	GAAGAGCAAGAGGCAGGCTCAGCAAAATGGTTACAGCCCAAGTCCC CGGTGGCTGT CAGTCA	60
409	QY	AAGCAAGCCCGGTTGTTATGACAAATGGAAAACACTATCAGATAAATCAACAGTGGGAGCG	468
61	Db	AAGCAAGCCCGGTTGTTATGACAAATGGAAAACACTATCAGATAAATCAACAGTGGGAGCG	120
469	QY	GACCTACCTAGGCAATCGGTTGGTTGTTACTGCTTATGGAGGAAGCCGAGGTTTTAACTG	528
121	Db	GACCTACCTAGGTAATGTGTGGTTGTTACTGCTTATGGAGGAAGCCGAGGTTTTAACTG	180
529	QY	CGAAAGTAAACCTGAAGCTGAAGAGACTGTGCTTTGACAAGTACACTGGGAACACTTACCG	588
181	Db	CGAAAGTAAACCTGAAGCTGAAGAGACTGTGCTTTGACAAGTACACTGGGAACACTTACCG	240
589	QY	AGTGGGTGACACTTATGAGCGGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCAATCGG	648
241	Db	AGTGGGTGACACTTATGAGCGGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCAATCGG	300
649	QY	GGCTGGGCGGAGGAGAAATAGCTGTACCATCGAAACCGCTGCCATGAAGGGGGTCAAGTC	708
301	Db	GGCTGGGCGGAGGAGAAATAGCTGTACCATCGAAACCGCTGCCATGAAGGGGGTCAAGTC	360
709	QY	CTACAAGATTGGTGACACCTGGAGGAGACCACATGAGACTGGTGGTTACATGTTTAGAGTG	768

Db 361 CTACAAAGATTGGTGAACCTGGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 420  
 769 TGTGTGCTTTGGTAAATGGAAAAGAGAAATGGAAGCTGCAAGCCCATAGCTGAGAAAGTGT 828  
 Db 421 TGTGTGCTTTGGTAAATGGAAAAGAGAAATGGAAGCTGCAAGCCCATAGCTGAGAAAGTGT 480  
 829 TGATCATGCTGCTGGGACTTCTATGTTGGTGGAGAAACCTGGGAGAAAGCCCTACCAAGG 888  
 Db 481 TGATCATGCTGCTGGGACTTCTATGTTGGTGGAGAAACCTGGGAGAAAGCCCTACCAAGG 540  
 889 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948  
 Db 541 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 949 TAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008  
 Db 601 TAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 1009 CAAGAAGGATAATGAGGAGAACTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 1068  
 Db 661 CAAGAAGGATAATGAGGAGAACTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 720  
 1069 GTGGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCGAGGGATCTGGCCCTTCAC 1128  
 Db 721 GTGGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCGAGGGATCTGGCCCTTCAC 780  
 1129 CGATGTTGCTGAGCTGTTTACCAACCGGAGCTTCAACCCAGGCTCTCCCTATGGCCCA 1188  
 Db 781 CGATGTTGCTGAGCTGTTTACCAACCGGAGCTTCAACCCAGGCTCTCCCTATGGCCCA 840  
 1189 CTGTGTACAGACAGTGGTGTGCTTACTCTGTGGGATGAGTGGCTGAAGACACAGG 1248  
 Db 841 CTGTGTACAGACAGTGGTGTGCTTACTCTGTGGGATGAGTGGCTGAAGACACAGG 900  
 1249 AAATAAGCAATGCTTTGACAGCTGCTGGGCAACGGAGTCACTGCTGCAAGAGACAGCTGT 1308  
 Db 901 AAATAAGCAATGCTTTGACAGCTGCTGGGCAACGGAGTCACTGCTGCAAGAGACAGCTGT 960  
 1309 AACCCAGACTTACGGTGGGAACTCAATGAGAGGCAATGCTGTTACCATTCACCTACAA 1368  
 Db 961 AACCCAGACTTACGGTGGGAACTTAAATGAGAGGCAATGCTGTTACCATTCACCTACAA 1020  
 1369 TGGCAGAGCTTCTACTCTGACACCAAGAGGCGACAGGACATCTTTGGTGGCAG 1428  
 Db 1021 TGGCAGAGCTTCTACTCTGACACCAAGAGGCGACAGGACATCTTTGGTGGCAG 1080  
 1429 CACAACCTTGAATTTATGAGCAGGACCAAGAAATCTCTTTCTGCACAGACCACTGTTT 1488  
 Db 1081 CACAACCTTGAATTTATGAGCAGGACCAAGAAATCTCTTTCTGCACAGACCACTGTTT 1140  
 1489 GGTTCAGACTCGAGAGGAAATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1548  
 Db 1141 GGTTCAGACTCGAGAGGAAATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 1549 CAACCAAAATTTACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608  
 Db 1201 CAACCAAAATTTACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 1609 GACACACAGAACTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1668  
 Db 1261 GACACACAGAACTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 1669 GGAATCTGCAACCAATGAGGGGTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728  
 Db 1321 GGAATCTGCAACCAATGAGGGGTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 1729 GCATGACATGAGGTCATGATGAGGTCATGATGAGGTCATGATGAGGTCATGATGAGGTCAT 1788  
 Db 1381 GCATGACATGAGGTCATGATGAGGTCATGATGAGGTCATGATGAGGTCATGATGAGGTCAT 1440  
 1789 ATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1848

Db 1441 ATGATTCGCTTACTCGCAACTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAATGT 1500  
 QY 1849 GAAACGACACATTCACAAAGCGTCTGAAAGAGGGGCAATGCTGAACTGTACATGCTTCGG 1908  
 Db 1501 GAAACGACACATTCACAAAGCGTCTGAAAGAGGGGCAATGCTGAACTGTACATGCTTCGG 1560  
 QY 1909 TCAGGGTGGGGCAGGTGGAAGTGTGATCCCGTGCACCAATGCCAGGATTCAGAGACTGG 1968  
 Db 1561 TCAGGGTGGGGCAGGTGGAAGTGTGATCCCGTGCACCAATGCCAGGATTCAGAGACTGG 1620  
 QY 1969 GAGCTTTTATCAAAATGGAGATTCATGAGGAGAGATGATGATGATGATGATGATGATGATG 2028  
 Db 1621 GAGCTTTTATCAAAATGGAGATTCATGAGGAGAGATGATGATGATGATGATGATGATGATG 1680  
 QY 2029 CTACTGCTATGSCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088  
 Db 1681 CTACTGCTATGSCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 QY 2089 CTCAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148  
 Db 1741 CTCAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
 QY 2149 CATCAGTGGATGACACACAGCCATCTCAATTTCCAAAGTACATTTCCAGTGGAGACC 2208  
 Db 1801 CATCAGTGGATGACACACAGCCATCTCAATTTCCAAAGTACATTTCCAGTGGAGACC 1860  
 QY 2209 TAAAAATTTCTGAGCCGTTGGAAGAGCTTACCATACAGGCCACTTTAAACTCTCTACAC 2268  
 Db 1861 TAAAAATTTCTGAGCCGTTGGAAGAGCTTACCATACAGGCCACTTTAAACTCTCTACAC 1920  
 QY 2269 CATCAAGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328  
 Db 1921 CATCAAGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
 QY 2329 CGGCCACCAAGAGTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388  
 Db 1981 CGGCCACCAAGAGTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
 QY 2389 CAGCAACACCTGACAGAGAGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2448  
 Db 2041 CAGCAACACCTGACAGAGAGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100  
 QY 2449 TGTGACCCGAAATTCAGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2508  
 Db 2101 TGTGACCCGAAATTCAGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
 QY 2509 GTGCGGATTCGCGGTGGAATATGAGCTGAGTGAAGGAGGATGAGGACAGTACCTGGA 2568  
 Db 2161 GTGCGGATTCGCGGTGGAATATGAGCTGAGTGAAGGAGGATGAGGACAGTACCTGGA 2220  
 QY 2569 TCTTCCAAAGCAGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2628  
 Db 2221 TCTTCCAAAGCAGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
 QY 2629 TGTAAATGTCTATCAGATATCTGAGATGAGGAGCAGATTTGATCTCTGCTCTACTTCA 2688  
 Db 2281 TGTAAATGTCTATCAGATATCTGAGATGAGGAGCAGATTTGATCTCTGCTCTACTTCA 2340  
 QY 2689 RACAAAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2748  
 Db 2341 AACAAAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
 QY 2749 TGTGTTGCTGCTGAGCAGACCCAGGCTCCCATCAAGGGTACAGAAATGATCTATTCGCC 2808  
 Db 2401 TGTGTTGCTGCTGAGCAGACCCAGGCTCCCATCAAGGGTACAGAAATGATCTATTCGCC 2460  
 QY 2809 ATCAGTGAAGGTAGCAGACAGBACTCAACCTTCCGAACTCGAACTCCGTCACCCCT 2868  
 Db 2461 ATCAGTGAAGGTAGCAGACAGBACTCAACCTTCCGAACTCGAACTCCGTCACCCCT 2520  
 QY 2869 CAGTGAATTCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2928  
 Db 2521 CAGTGAATTCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580









QY 7309 GGTGACAGAGAGATTCCTGAGAGATAATCATCTTTCCAAATCCAGAGAAACAAGCATGT 7368  
DB 6961 GGTGACAGAGAGATTCCTGAGAGATAATCATCTTTCCAAATCCAGAGAAACAAGCATGT 7020  
QY 7369 CTCCTGCGCAAGATCATCTAACTGGAGTGATTTAGCAGACCCAGCTTAGAGTTCTTC 7428  
DB 7021 CTCCTGCGCAAGATCATCTAACTGGAGTGATTTAGCAGACCCAGCTTAGAGTTCTTC 7080  
QY 7429 TTTCTTTCTTAAGCCCTTTGCTCTCGAGAGAGATTTCTCCAGCTTCAGCTCAAATCAGCT 7488  
DB 7081 TTTCTTTCTTAAGCCCTTTGCTCTCGAGAGAGATTTCTCCAGCTTCAGCTCAAATCAGCT 7140  
QY 7489 TCTCCAAGCATCACCTCGGAGTTTCTCGAGAGATTTCTCAATAATGAGGCGTCCACATT 7548  
DB 7141 TCTCCAAGCATCACCTCGGAGTTTCTCGAGAGATTTCTCAATAATGAGGCGTCCACATT 7200  
QY 7549 GCTGTCTGCTTCCGAAGTATTAATACCGCTCAGTATTTTAAATGAAGTCAATTAAGA 7608  
DB 7201 GCTGTCTGCTTCCGAAGTATTAATACCGCTCAGTATTTTAAATGAAGTCAATTAAGA 7260  
QY 7609 TTTGTTTGGGATCAATAGGAAGCATATGCGACCAACCAAGATGCAATGTTTGAAT 7668  
DB 7261 TTTGTTTGGGATCAATAGGAAGCATATGCGACCAACCAAGATGCAATGTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAGTAGGAAGTCAACCCAAACACCTTCTGCTTTCACTTAAGTGT 7728  
DB 7321 GATATGACCAAAATTTTAAGTAGGAAGTCAACCCAAACACCTTCTGCTTTCACTTAAGTGT 7380  
QY 7729 CTGGCCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7788  
DB 7381 CTGGCCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7440  
QY 7789 CAGTACT 7795  
DB 7441 CAGTACT 7447

RESULT 7  
5455158-2  
; Patent No. 5455158  
; APPLICANT: VOGEL, TIKVA, LEVANON, AVIGDOR, WERBER, MOSHE M.;  
; GUY, RACHEL; PANET, AMOS  
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND  
; USES AND METHODS OF PRODUCING SAME  
; NUMBER OF SEQUENCES: 20  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/05/58,241  
; FILING DATE: 04-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 526,397  
; FILING DATE: 21-MAY-1990  
; APPLICATION NUMBER: 345,952  
; FILING DATE: 28-APR-1989  
; APPLICATION NUMBER: 291,951  
; FILING DATE: 29-DEC-1988  
; SEQ ID NO: 2:  
; LENGTH: 7705  
5455158-2

Query Match 95.1%; Score 7413.4; DB 6; Length 7705;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 7426; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 349 GAAGACAGAGAGCGGCTCAGCAATGGTTTCAGCCCGGCTCCCGGTGGCTGTCACTCA 408  
DB 1 GAAGACAGAGAGCGGCTCAGCAATGGTTTCAGCCCGGCTCCCGGTGGCTGTCACTCA 60  
QY 409 AAGCAAGCCCGGTGTGTATGACAATGGAAAACTATCAGATAAATCAACAGTGGAGCG 468  
DB 61 AAGCAAGCCCGGTGTGTATGACAATGGAAAACTATCAGATAAATCAACAGTGGAGCG 120  
QY 469 GACCTACCTAGGCAATGCGTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTG 528

DB 121 GACCTACCTAGGCAATGCTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTG 180  
QY 529 CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTCTTTGCAAGTAGTACACTTGGGAACACTTACCG 588  
DB 181 CGAAAGTAAACCTGAAAGCTGAAGAGACTTGTCTTTGCAAGTAGTACACTTGGGAACACTTACCG 240  
QY 589 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACCTGATCGG 648  
DB 241 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACCTGATCGG 300  
QY 649 GGTGCGCGAGGAGAGATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCACTC 708  
DB 301 GGTGCGCGAGGAGAGATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCACTC 360  
QY 709 CTACAAGATGTTGACACCTCGAGGAGACCAATGAGACTGCTGTTTACATGTAGAGTG 768  
DB 361 CTACAAGATGTTGACACCTCGAGGAGACCAATGAGACTGCTGTTTACATGTAGAGTG 420  
QY 769 TGTGTCTTGTGTAATGGAAGAGAGATGGAACCTGCAAGCCCATAGCTGAGAAGTGT 828  
DB 421 TGTGTCTTGTGTAATGGAAGAGAGATGGAACCTGCAAGCCCATAGCTGAGAAGTGT 480  
QY 829 TGATCATGCTGTGGAGACTTCTATGTTGTCGAGAAACGTGGGAGAACCCCTACCAAG 888  
DB 481 TGATCATGCTGTGGAGACTTCTCTATGTTGTCGAGAAACGTGGGAGAACCCCTACCAAG 540  
QY 889 CTGATGATGTTGTAATGGAAGAGAGATGGAACCTGCAAGCCCATAGCTGAGAAGTGT 948  
DB 541 CTGATGATGTTGTAATGGAAGAGAGATGGAACCTGCAAGCCCATAGCTGAGAAGTGT 600  
QY 949 TAGAAATAGATGCAACGATCAGGACACAAAGACATCTTATAGAAATGAGAGACACTTGGAG 1008  
DB 601 TAGAAATAGATGCAACGATCAGGACACAAAGACATCTTATAGAAATGAGAGACACTTGGAG 660  
QY 1009 CAAGAAAGATATCGAGGAGAAACCTGCTCCAGTGATCTGCAAGGCAACGCGGAGGAGA 1068  
DB 661 CAAGAAAGATATCGAGGAGAAACCTGCTCCAGTGATCTGCAAGGCAACGCGGAGGAGA 720  
QY 1069 GTGGAAGTGTGAGAGGACACCTTGTGAGACCAATCGAGCGGATCTGGCCCTTTAC 1128  
DB 721 GTGGAAGTGTGAGAGGACACCTTGTGAGACCAATCGAGCGGATCTGGCCCTTTAC 780  
QY 1129 CGATGTTGTCGAGCTGTTTACCAACCGCAGCTCACCCCGAGCTTCCCTCTATGGCCA 1188  
DB 781 CGATGTTGTCGAGCTGTTTACCAACCGCAGCTCACCCCGAGCTTCCCTCTATGGCCA 840  
QY 1189 CTGTGTACAGACAGTGTGTGTCTTACTCTTGTGGGATGCACTGTGAGACACAAAG 1248  
DB 841 CTGTGTACAGACAGTGTGTGTCTTACTCTTGTGGGATGCACTGTGAGACACAAAG 900  
QY 1249 AATAAGCAATGCTTTGACAGCTGCTGGCAACGAGTCACTGCCAGAGACAGCTGT 1308  
DB 901 AATAAGCAATGCTTTGACAGCTGCTGGCAACGAGTCACTGCCAGAGACAGCTGT 960  
QY 1309 AACCAGACTTACGCTGGCAACCTCAAAATGAGAGCCATGTGTCTTACCATTCACCTACAA 1368  
DB 961 AACCAGACTTACGCTGGCAACCTCAAAATGAGAGCCATGTGTCTTACCATTCACCTACAA 1020  
QY 1369 TGGCAGAGACTTCTACTCTGCAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1428  
DB 1021 TGGCAGAGACTTCTACTCTGCAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
QY 1429 CACAATCTCGAATATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1488  
DB 1081 CACAATCTCGAATATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1489 GGTTCAGACTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1548  
DB 1141 GGTTCAGACTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1549 CAACCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1608





b	5581	TGACAATGCTCGGAGCTCCCTCTGTGGTTCATCGACGCTTCCACTGCCAATGTGATGCAACATC	5644
y	5589	CAACCTGCGTTTCCTGGCCACCAACACCCAAATTCCTTGTGTATCATGCGAGCGGCCACG	6048
b	5641	CAACCTGCGTTTCCTGGCCACCAACACCCAAATTCCTTGTGTATCATGCGAGCGGCCACG	5700
y	6049	TGCCAGGATTACCGGCTACATCATCAAGTATGAGAAGCCTGGTCTCTCTCCGAGAGAGT	6108
b	5701	TGCGAGGATTACCGGCTACATCATCAAGTATGAGAAGCCTGGTCTCTCTCCGAGAGAGT	5760
y	6109	GGTCCCTCGGCCCCCGCTGTGTACAGAGGCTACTATTACTGGCTCGAAACCGGGAAC	6168
b	5761	GGTCCCTCGGCCCCCGCTGTGTACAGAGGCTACTATTACTGGCTCGAAACCGGGAAC	5820
y	6169	CGAATATACAAATTATGTCAATTGCCCTGAAGAATAATCAGAAGAGCGAGCCCTGATTTGG	6228
b	5821	CGAATATACAAATTATGTCAATTGCCCTGAAGAATAATCAGAAGAGCGAGCCCTGATTTGG	5880
y	6229	AAGGAAAAAGACAGACGAGCTTCCCAACTGTGTACCTTCCACACCCCAATCTTCATGG	6288
b	5881	AAGGAAAAAGACAGACGAGCTTCCCAACTGTGTACCTTCCACACCCCAATCTTCATGG	5940
y	6289	ACCAGAGATCTTGGATGTTCTTTCACAGTTCAAAAGACCCCTTTCTGTCAACCAACCTTGG	6348
b	5941	ACCAGAGATCTTGGATGTTCTTTCACAGTTCAAAAGACCCCTTTCTGTCAACCAACCTTGG	6000
y	6349	GTATGACACTGGAAATGGTATTTCAGCTTCCCTGGCACCTTCTGGTCAGCAACCCAGTGTGG	6408
b	6001	GTATGACACTGGAAATGGTATTTCAGCTTCCCTGGCACCTTCTGGTCAGCAACCCAGTGTGG	6060
y	6409	GCAACAAATGATCTTTGAGGAAACATGGTTTTAGCGGACCAACACCGGCCCAACAGGCCAC	6468
b	6061	GCAACAAATGATCTTTGAGGAAACATGGTTTTAGCGGACCAACACCGGCCCAACAGGCCAC	6120
y	6469	CCCCATAGGCCATAGGCCAAGACCAATACCCGCCGAATGTAGGACAAAGAGCTCTCTCTCA	5528
b	6121	CCCCATAGGCCATAGGCCAAGACCAATACCCGCCGAATGTAGGACAAAGAGCTCTCTCTCA	6180
y	6529	GACAAACATCTCTATGGGCCCCCATTCAGGACACTTCTCAGTATCATCATTTTCACTGTCTATCC	6588
b	6181	GACAAACATCTCTATGGGCCCCCATTCAGGACACTTCTCAGTATCATCATTTTCACTGTCTATCC	6240
y	6589	TGTTGGCACTGATGAAGAACCCCTTACGTTACAGTTACAGTTCTCTACCAAGTGTGCCAC	6648
b	6241	TGTTGGCACTGATGAAGAACCCCTTACGTTACAGTTACAGTTCTCTACCAAGTGTGCCAC	6300
y	6649	TCTGACAGGCTTCAACAGAGGTGCCACCTTACAAACATCATAGTGAGAGCACTGGAAGACCA	6708
b	6301	TCTGACAGGCTTCAACAGAGGTGCCACCTTACAAACATCATAGTGAGAGCACTGGAAGACCA	6360
y	6709	GCAGAGSCATAAGGTTTCGGGAAAGAGTGTGTTACCGTGGGCAACTCTGTCAACGGAAGCTT	6768
b	6361	GCAGAGSCATAAGGTTTCGGGAAAGAGTGTGTTACCGTGGGCAACTCTGTCAACGGAAGCTT	6420
y	6769	GAACCACTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCCGTGG	6828
b	6421	GAACCACTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCCGTGG	6480
y	6829	AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTTAAAGCTGTTGTGCCAGTGTCTTAGGCTT	6888
b	6481	AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTTAAAGCTGTTGTGCCAGTGTCTTAGGCTT	6540
y	6889	TGGAAGTGGTTCATTTTCAGATGTGATTTCATCTAGATGGTGGCCATGACAAATGGTGTGAACCTA	6948
b	6541	TGGAAGTGGTTCATTTTCAGATGTGATTTCATCTAGATGGTGGCCATGACAAATGGTGTGAACCTA	6600
y	6949	CAAGATTGGAGAGAAATGGGACCGTTCAGGGGAGAAATGGCCAGATGATGAGCTGCACATG	7008
b	6601	CAAGATTGGAGAGAAATGGGACCGTTCAGGGGAGAAATGGCCAGATGATGAGCTGCACATG	6660
y	7009	TCTTGGAAACGGAAAAGGAAATTCAAAGTGTGACCCCTCATGAGGCAACCGTGTTCACGATGA	7068
b	6661	TCTTGGAAACGGAAAAGGAAATTCAAAGTGTGACCCCTCATGAGGCAACCGTGTTCACGATGA	6720

Qy	7069	TGGGAGAGACATACCA	CGTAGGAGAA	CAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC	71228
Db	6721	TGGGAGAGACATACCA	CGTAGGAGAA	CAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC	6780
Qy	7129	CTGCACATGCTTTGGAGGCCAGCGGGCTCTGACAACTGCGCAGACCTGGGG	7188		
Db	6781	CTGCACATGCTTTGGAGGCCAGCGGGCTCTGACAACTGCGCAGACCTGGGG	6840		
Qy	7189	TGAACCCAGTCCCGAAGGCATCTATGCGCAGTCCCTACAAACCAGTATTTCTCAGAGATACCA	7248		
Db	6841	TGAACCCAGTCCCGAAGGCATCTATGCGCAGTCCCTACAAACCAGTATTTCTCAGAGATACCA	6900		
Qy	7249	TCAGAGAAACAAACATTAATGTGTTAAATTCGCCAAATGAGTGTCTCATGCCCTTTAGATGTACA	7308		
Db	6901	TCAGAGAAACAAACATTAATGTGTTAAATTCGCCAAATGAGTGTCTCATGCCCTTTAGATGTACA	6960		
Qy	7309	GGCTGACAGAGAAGATTCCGAGAGTAATATCATCTTTCCAAATCCAGAGGAACAAGCATGT	7368		
Db	6961	GGCTGACAGAGAAGATTCCGAGAGTAATATCATCTTTCCAAATCCAGAGGAACAAGCATGT	7020		
Qy	7369	CTCTCTGCCAAGATCCATCTAAATCGAGTGATGTTTAGCAGACCCAGCTTAGAGTTCTTC	7428		
Db	7021	CTCTCTGCCAAGATCCATCTAAATCGAGTGATGTTTAGCAGACCCAGCTTAGAGTTCTTC	7080		
Qy	7429	TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTTCCAGGTTTCAGCTCAACTCACAGCT	7488		
Db	7081	TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTTCCAGGTTTTCAGCTCAACTCACAGCT	7140		
Qy	7489	TCCTCAAGCATCACCTCGGGAGTTTCTCGAGGGTTTCTCAATAATGAGGGCTGACACATT	7548		
Db	7141	TCCTCAAGCATCACCTCGGGAGTTTCTCGAGGGTTTCTCAATAATGAGGGCTGACACATT	7200		
Qy	7549	GCCTGTTCTGCTTCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7608		
Db	7201	GCCTGTTCTGCTTCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7260		
Qy	7609	TTTGGTTTGGATCAATAGGAAGCATATGACGCCAAACCAAGATGCAAAATGTTTTCAAAT	7668		
Db	7261	TTTGGTTTGGATCAATAGGAAGCATATGACGCCAAACCAAGATGCAAAATGTTTTCAAAT	7320		
Qy	7669	GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAGTGT	7728		
Db	7321	GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAGTGT	7380		
Qy	7729	CTGGCCCGCAATCTCTAGGAAACAAGCATGATCTTGTTACTGTGATATTTTAAATATCCA	7788		
Db	7381	CTGGCCCGCAATCTCTAGGAAACAAGCATGATCTTGTTACTGTGATATTTTAAATATCCA	7440		
Qy	7789	CAGTACT	7795		
Db	7441	CAGTACT	7447		

```

RESULT 8
US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: ZymGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```









QY 4064 ----- 4063  
DB GCATCAGTAGTGGCGCAGGAGAGGTATCCCTATTTTGAAGATTTGTGTACTCCT 3960  
QY 4064 ----- 4063  
DB CAGTAGGATACTACACAGTACAGGGCTGGAGCGGGCAATTGACTATGATATCAGCGTTA 4020  
QY 4064 ----- 4070  
DB TCACTCTCAATTAATGGCGCGAGAGTGCCTCTATACACTGACACAAACAGCGCTGTT 4080  
QY 4071 CTCCTCCCACTGACCTGGATTCACCAACATTTGGTCCAGACACCATCGGTGTCACTGGG 4130  
DB CTCCTCCCACTGACCTGGATTCACCAACATTTGGTCCAGACACCATCGGTGTCACTGGG 4140  
QY 4131 CTCACCCCCATTCATTTGATTTAACCAACTCTCTGGTGGTGTACTCACTCTGTGAAAATG 4190  
DB CTCACCCCCATTCATTTGATTTAACCAACTCTCTGGTGGTGTACTCACTCTGTGAAAATG 4200  
QY 4191 AGGAAGATGTTGCAGAGTGTCAATTTCTCCTTCAGACAAATGCAAGTGGTCTTTAAACAATC 4250  
DB AGGAAGATGTTGCAGAGTGTCAATTTCTCCTTCAGACAAATGCAAGTGGTCTTTAAACAATC 4260  
QY 4251 TCCTGCCGTGACAGAAATATGTATGTATGTCTCTCAGTGTCTTACGACAAATGCAAGTGGT 4310  
DB TCCTGCCGTGACAGAAATATGTATGTATGTCTCTCAGTGTCTTACGACAAATGCAAGTGGT 4320  
QY 4311 CACCTCTTAGAGGAGACAGAAAACAGGTCTTTGATTTCCCACTGGCATTTGATTTTCTG 4370  
DB CACCTCTTAGAGGAGACAGAAAACAGGTCTTTGATTTCCCACTGGCATTTGATTTTCTG 4380  
QY 4371 ATATTACTGCCAATCTTTTACTGTGCTAGTGGTCTCTCGAGCCCAATCTAGCTGCT 4430  
DB ATATTACTGCCAATCTTTTACTGTGCTAGTGGTCTCTCGAGCCCAATCTAGCTGCT 4440  
QY 4431 ACAGATCGGCATCATCCCGAGACTTCAGTGGGAGACCTCGAGAGATCGGGTGGCCCC 4490  
DB ACAGATCGGCATCATCCCGAGACTTCAGTGGGAGACCTCGAGAGATCGGGTGGCCCC 4500  
QY 4491 ACTCTCGGAATTCATACCCCTCACCAACCTCACTCCAGGACAGAGATGTGTGTACGA 4550  
DB ACTCTCGGAATTCATACCCCTCACCAACCTCACTCCAGGACAGAGATGTGTGTGTACGA 4560  
QY 4551 TCGTTGCTTTAATGGCAGAGAGAAATCCCTTATGATTTGGCCCAACATCAACAGTTT 4610  
DB TCGTTGCTTTAATGGCAGAGAGAAATCCCTTATGATTTGGCCCAACATCAACAGTTT 4620  
QY 4611 CTGATGTTCCGAGGACCTGGAAATTTGTCTGCGACCCCAACAGCTCTACTGATCAGCT 4670  
DB CTGATGTTCCGAGGACCTGGAAATTTGTCTGCGACCCCAACAGCTCTACTGATCAGCT 4680  
QY 4671 GGGATGCTCTGCTGACAGTGAGATATTAAGAGTCACTTACGGAGAAACAGAGGAA 4730  
DB GGGATGCTCTGCTGACAGTGAGATATTAAGAGTCACTTACGGAGAAACAGAGGAA 4740  
QY 4731 ATAGCCCTGTCCAGAGTTCACTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGGCC 4790  
DB ATAGCCCTGTCCAGAGTTCACTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGGCC 4800  
QY 4791 TTAACCTGGAGTTGATTTATAACAATCACTGTGTATGTCTGCTGCTGCTGCTGCTGCTGCT 4850  
DB TTAACCTGGAGTTGATTTATACCATCACTGTGTATGTCTGCTGCTGCTGCTGCTGCTGCT 4860  
QY 4851 CGCAAGCAGCAGCAAGCAATTTCCATTAATTCAGAACAGAAATTCAGAAACCATCCAGA 4910  
DB CGCAAGCAGCAGCAAGCAATTTCCATTAATTCAGAACAGAAATTCAGAAACCATCCAGA 4920  
QY 4911 TGCAAGTGAACAGATGTTTACAGCAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCC 4970  
DB TGCAAGTGAACAGATGTTTACAGCAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCC 4980  
QY 4971 CTGTTACTGGTTACAGAGTAAACCACTCCCAAAAATGAGCAGGACCAACAAAACCTA 5030

DB 4981 CTGTTACTGGTTACAGAGTAAACCACTCCCAAAAATGAGCAGGACCAACAAAACCTA 5040  
QY 5031 AAATCGAGGTCAGATCAAAACAGAAATGACTATTTGAAGGCTTGAGCCCAAGTGGAGT 5090  
DB 5041 AAATCGAGGTCAGATCAAAACAGAAATGACTATTTGAAGGCTTGAGCCCAAGTGGAGT 5100  
QY 5091 ATGTGTTAGTGTCTATGCTCAGAAATCAAGCGGAGAGAGTCAAGCTCTGGTTCAAGCTG 5150  
DB 5101 ATGTGTTAGTGTCTATGCTCAGAAATCAAGCGGAGAGAGTCAAGCTCTGGTTCAAGCTG 5160  
QY 5151 CAGTAAACCAAAATGATCGCCCTAAAGGATGCGATTCATCTGATGTGATGTCGATTTCCA 5210  
DB 5161 CAGTAAACCAAAATGATCGCCCTAAAGGATGCGATTCATCTGATGTGATGTCGATTTCCA 5220  
QY 5211 TCAAAATTCCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTCGA 5270  
DB 5221 TCAAAATTCCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGAACCTACTCGA 5280  
QY 5271 GGCCTGAGGATGGAATCCATGAGCTATTTCCCTGCACTTCCAGGTGAAGAGACACTGCG 5330  
DB 5281 GGCCTGAGGATGGAATCCATGAGCTATTTCCCTGCACTTCCAGGTGAAGAGACACTGCG 5340  
QY 5331 AGCTGCAAGGCTCAGACCGGCTCTGAGTACACAGTCAAGTGTGCTTGCCTTGCACGATG 5390  
DB 5341 AGCTGCAAGGCTCAGACCGGCTCTGAGTACACAGTCAAGTGTGCTTGCCTTGCACGATG 5400  
QY 5391 ATATGAGAGCCAGGCCCTGATTTGAAACCAAGTCCACAGTATTTCTGCAACCACTGACC 5450  
DB 5401 ATATGAGAGCCAGGCCCTGATTTGAAACCAAGTCCACAGTATTTCTGCAACCACTGACC 5460  
QY 5451 TGAAGTTCACTCAGGTTCACACCAAGCTCAGCGCCAGTGGACACCAACCAATGTTT 5510  
DB 5461 TGAAGTTCACTCAGGTTCACACCAAGCTCAGCGCCAGTGGACACCAACCAATGTTT 5520  
QY 5511 AGCTCACTGATATCGAGTGGGCTGACCCCAAGGAGAGACCGGACCAATGAAGAAA 5570  
DB 5521 AGCTCACTGATATCGAGTGGGCTGACCCCAAGGAGAGACCGGACCAATGAAGAAA 5580  
QY 5571 TCAACCTTCTCTGACAGTCACTCGGTGTTGATCAGGACTTATGTTGGCCACCAAT 5630  
DB 5581 TCAACCTTCTCTGACAGTCACTCGGTGTTGATCAGGACTTATGTTGGCCACCAAT 5640  
QY 5631 ATGAAGTGTGTCTATGCTCTTAAGGACCACTTTGACAGAGACCAAGCTCAGGGTGTG 5690  
DB 5641 ATGAAGTGTGTCTATGCTCTTAAGGACCACTTTGACAGAGACCAAGCTCAGGGTGTG 5700  
QY 5691 TCAACCTCTCTGAGAGATGTCAGCCCAAGAGGCTCGTGTGACAGATGCTACTGAGA 5750  
DB 5701 TCAACCTCTCTGAGAGATGTCAGCCCAAGAGGCTCGTGTGACAGATGCTACTGAGA 5760  
QY 5751 CCACCATCACTTACTGAGAGACCAAGACTGAGACGATCACTGGCTTCCAAGTTGATG 5810  
DB 5761 CCACCATCACTTACTGAGAGACCAAGACTGAGACGATCACTGGCTTCCAAGTTGATG 5820  
QY 5811 CCCTTCCAGCAATGCGCAGTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAGCT 5870  
DB 5821 CCCTTCCAGCAATGCGCAGTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAGCT 5880  
QY 5871 ACACCATCACTGCTTACCAACAGGACCTGACTTACAGATCTACCTGTACACCTTGAATG 5930  
DB 5881 ACACCATCACTGCTTACCAACAGGACCTGACTTACAGATCTACCTGTACACCTTGAATG 5940  
QY 5931 ACAATGCTGGAGCTCCCTGTGTTGATCGAAGCTTCACTGGCCTTGTGATGACCACTCA 5990  
DB 5941 ACAATGCTGGAGCTCCCTGTGTTGATCGAAGCTTCACTGGCCTTGTGATGACCACTCA 6000  
QY 5991 ACCTGGCTTCTTGGCCACCAACCAATTCCTTGTGTTGATCATGGCAGCCGACCTG 6050  
DB 6001 ACCTGGCTTCTTGGCCACCAACCAATTCCTTGTGTTGATCATGGCAGCCGACCTG 6060  
QY 6051 CCAGGATTAACGGCTACATCATCAAGTATGAGAGGCTGGTCTCTCCAGAGAGATGG 6110

6061 CCAGGATTACCGGCTACATCATCAAGTATGAGAAGCCTGGGTCTCTCCCGAGAGAAGTGG 6120  
6111 TCCCTCGGCCCCCGCTGGTGTACAGAGGCTACTATTACTGCGCTTGGAAACCGGAAACCG 6170  
6121 TCCCTCGGCCCCCGCTGGTGTACAGAGGCTACTATTACTGCGCTTGGAAACCGGAAACCG 6180  
6171 AATATACAAATTTATGTCTATTCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 6230  
6181 AATATACAAATTTATGTCTATTCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 6240  
6231 GGAAG 6290  
6241 GGAAG 6300  
6291 CAGAGATCTTGGATGTTCTTCCACAGTCAAAAGAGCCCTTTCGTCACCCACCTGGGT 6350  
6301 CAGAGATCTTGGATGTTCTTCCACAGTCAAAAGAGCCCTTTCGTCACCCACCTGGGT 6360  
6351 ATGACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6410  
6361 ATGACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6420  
6411 AACAAATGATCTTGGAG 6470  
6421 AACAAATGATCTTGGAG 6480  
6471 CCATAAGGCTATGAGGAG 6530  
6481 CCATAAGGCTATGAGGAG 6540  
6531 CAACATCTCATGCGGCGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6590  
6541 CAACATCTCATGCGGCGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6600  
6591 TTGCACTGATGAAG 6650  
6601 TTGCACTGATGAAG 6660  
6651 TGACAGGCTTACAG 6710  
6661 TGACAGGCTTACAG 6720  
6711 AGAGGCTATAGGTTGAG 6770  
6721 AGAGGCTATAGGTTGAG 6780  
6771 ACCAAGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6830  
6781 ACCAAGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6840  
6831 ATGAGTGGAG 6890  
6841 ATGAGTGGAG 6900  
6891 GAAGTGGTCAATTTAG 6950  
6901 GAAGTGGTCAATTTAG 6960  
6951 AGATTGAG 7010  
6961 AGATTGAG 7020  
7011 TTGGAG 7070  
7021 TTGGAG 7080  
7071 GGAAGAGATACCACTGAG 7130  
7081 GGAAGAGATACCACTGAG 7140  
7131 GCACATGCTTTGGAG 7190  
7141 GCACATGCTTTGGAG 7200

QY 7191 AACCCAGTCCCGAGGCACTACTGCCAGTCTTACAAACAGTATTTCTCAGAGATACCATC 7250  
DB 7201 AACCCAGTCCCGAGGCACTACTGCCAGTCTTACAAACAGTATTTCTCAGAGATACCATC 7260  
QY 7251 AGAGAAACAAACACATAATGTTAAATTTGCCCAATTTAGTGTCTTCAATCCAGAGGAAAGCATGCT 7310  
DB 7261 AGAGAAACAAACACATAATGTTAAATTTGCCCAATTTAGTGTCTTCAATCCAGAGGAAAGCATGCT 7320  
QY 7311 CTGACAGAGAGATCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAAGCATGCT 7370  
DB 7321 CTGACAGAGAGATCCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAAGCATGCT 7380  
QY 7371 CTCTCCCAAGATCCATCTAAACTGAGTGTGTTAGCAGACCCAGCTTTAGAGTCTTCTT 7430  
DB 7381 CTCTCCCAAGATCCATCTAAACTGAGTGTGTTAGCAGACCCAGCTTTAGAGTCTTCTT 7440  
QY 7431 TCTTTCTTTAAGCCCTTTGCTCTGAGGAGTCTTCTGAGGAGTCTTCTCATAAATGAGGCTTGCATTC 7490  
DB 7441 TCTTTCTTTAAGCCCTTTGCTCTGAGGAGTCTTCTGAGGAGTCTTCTCATAAATGAGGCTTGCATTC 7500  
QY 7491 TCCAAAGCATCACCTTGGGAGTCTTCTGAGGAGTCTTCTCATAAATGAGGCTTGCATTC 7550  
DB 7501 TCCAAAGCATCACCTTGGGAGTCTTCTGAGGAGTCTTCTCATAAATGAGGCTTGCATTC 7560  
QY 7551 CTGTTCTGCTTGAAGTATCAATACCCCTCAGTATTTTAAATGAAGTATCTAAGATT 7610  
DB 7561 CTGTTCTGCTTGAAGTATCAATACCCCTCAGTATTTTAAATGAAGTATCTAAGATT 7620  
QY 7611 TGGTTTGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAAAATGTTTGAATGA 7670  
DB 7621 TGGTTTGGATCAATAGGAAAGCATATGACGCAACCAAGATGCAAAATGTTTGAATGA 7680  
QY 7671 TATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACACTTCTGCTTTTCAATTAAGTGTCT 7730  
DB 7681 TATGACCAAAATTTTAAAGTAGGAAAGTCAACCAACACTTCTGCTTTTCAATTAAGTGTCT 7740  
QY 7731 GCGCCGCAATCTGTAGGAAAGCATATGCTTGTACTGTATTTTAAATATCCACA 7790  
DB 7741 GCGCCGCAATCTGTAGGAAAGCATATGCTTGTACTGTATTTTAAATATCCACA 7800  
QY 7791 GTA 7793  
DB 7801 GTA 7803

## RESULT 9

PCT-US93-12687-1  
; Sequence 1, Application PC/TUS9312687  
; GENERAL INFORMATION:  
; APPLICANT: Irani, Meher H.  
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 4225 Roosevelt Way, N.E.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/12687  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/998,271  
; FILING DATE: 31-DEC-1992  
; ATTORNEY/AGENT INFORMATION:







2Y 4071 CTCCTCCACCTGACCTGGATTACCAACATTTGGTCCAGACACATGCGTGCACCTGG 4130  
2b CTCCTCCACCTGACCTGGATTACCAACATTTGGTCCAGACACATGCGTGCACCTGG 4140  
2Y 4131 CTCACCCCTCCATTCATTTAAACCAATTCCTGCTGGTGGTACTCAGCTGTGAAATG 4190  
2b CTCACCCCTCCATTCATTTAAACCAATTCCTGCTGGTGGTACTCAGCTGTGAAATG 4200  
2Y 4191 AGGAAGATGTTGAGAGTTGCAATTTCTCTTCAGACATGAGTGGTCTTAAACAAATC 4250  
2b AGGAAGATGTTGAGAGTTGCAATTTCTCTTCAGACATGAGTGGTCTTAAACAAATC 4260  
2Y 4251 TCCTGCTGGTACAGATATGATGAGTGTCTCCAGTGTCTACGAAACATCAGAGCA 4310  
2b TCCTGCTGGTACAGATATGATGAGTGTCTCCAGTGTCTACGAAACATCAGAGCA 4320  
2Y 4311 CACCTCTTAGAGGAAGACAGAAACAGGCTCTGATTCCTCCAACTGGGATTTGCTTCTG 4370  
2b CACCTCTTAGAGGAAGACAGAAACAGGCTCTGATTCCTCCAACTGGGATTTGCTTCTG 4380  
2Y 4371 ATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGGCT 4430  
2b ATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGGCT 4440  
2Y 4431 ACAGATCCGCCATCATCCGAGCACTTCAGTGGGAGACTTCAGAGAGATCGGGTGGCCC 4490  
2b ACAGATCCGCCATCATCCGAGCACTTCAGTGGGAGACTTCAGAGAGATCGGGTGGCCC 4500  
2Y 4491 ACTCTCGGAATTCATCACCCTCACCACCTCACTCCAGGACAGAGATGTTGGTCA 4550  
2b ACTCTCGGAATTCATCACCCTCACCACCTCACTCCAGGACAGAGATGTTGGTCA 4560  
2Y 4551 TCGTGTCTTTAATGGCAGAGAGAAAGTCCCTTATTGATGGCCAAACATCAACAGTTT 4610  
2b TCGTGTCTTTAATGGCAGAGAGAAAGTCCCTTATTGATGGCCAAACATCAACAGTTT 4620  
2Y 4611 CTGATGTTCCAGGAGCTGGAAGTGTGCTGGACCCCGCCAGCCTACTGATCAGCT 4670  
2b CTGATGTTCCAGGAGCTGGAAGTGTGCTGGACCCCGCCAGCCTACTGATCAGCT 4680  
2Y 4671 GGGATGCTCTGCTGTACAGTGGATATTAAGGATCACTTACGGAGAAACAGGAGAA 4730  
2b GGGATGCTCTGCTGTACAGTGGATATTAAGGATCACTTACGGAGAAACAGGAGAA 4740  
2Y 4731 ATAGCCCTGTCCAGAGTTCACTGCTGGAGCAAGTCTACAGTACCATCAGCGGCC 4790  
2b ATAGCCCTGTCCAGAGTTCACTGCTGGAGCAAGTCTACAGTACCATCAGCGGCC 4800  
2Y 4791 TTAACCTGGAGTTCATTTATACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAGCC 4850  
2b TTAACCTGGAGTTCATTTATACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAGCC 4860  
2Y 4851 CCGCAAGGACGAGCAATTTCCATTAATCCGACAGAAATGACAAACCATCCAGA 4910  
2b CCGCAAGGACGAGCAATTTCCATTAATCCGACAGAAATGACAAACCATCCAGA 4920  
2Y 4911 TCAGAGTACCATGTTTCAGGACAAACAGATTTAGTGTCAAGTGGTGTGCTTCAAGTTCCC 4970  
2b TCAGAGTACCATGTTTCAGGACAAACAGATTTAGTGTCAAGTGGTGTGCTTCAAGTTCCC 4980  
2Y 4971 CTGTATCTGGTTACAGAGTAAACCACTCCCAAAATGGAACAGGACCAACAAACTA 5030  
2b CTGTATCTGGTTACAGAGTAAACCACTCCCAAAATGGAACAGGACCAACAAACTA 5040  
2Y 5031 AAACCTGAGTCCAGATCAAAACAGAAATGACATTTGAAGGCTTTCAGGCCCCACAGTGGAGT 5090  
2b AAACCTGAGTCCAGATCAAAACAGAAATGACATTTGAAGGCTTTCAGGCCCCACAGTGGAGT 5100  
2Y 5091 ATGTGTTAGTGTCTATGCTCAGATTCCAAGCGGAGAGAGTCAAGCTCTGTTTCAAGCTG 5150  
2b ATGTGTTAGTGTCTATGCTCAGATTCCAAGCGGAGAGAGTCAAGCTCTGTTTCAAGCTG 5160

5151 CAGTAACCAACATTCATGCTGCGCTAAAGGATGCGCATTCACGTGATGTGATGTCCGATCCA 5210  
5161 CAGTAACCAACATTCATGCTGCGCTAAAGGATGCGCATTCACGTGATGTGATGTCCGATCCA 5220  
5211 TCAGAAATTCCTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGAGTACCTACTCGA 5270  
5221 TCAGAAATTCCTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGAGTACCTACTCGA 5280  
5271 GCGCTGAGAGTGAATCCATGAGCTATTCCCTGCACCTGATGGTGAAGAGACACTGCAG 5330  
5281 GCGCTGAGAGTGAATCCATGAGCTATTCCCTGCACCTGATGGTGAAGAGACACTGCAG 5340  
5331 AGCTGCAAGGCTCAGACCGGCTTCTGAGTACACAGTCACTGAGTGTGTTGCTTGCACGATG 5390  
5341 AGCTGCAAGGCTCAGACCGGCTTCTGAGTACACAGTCACTGAGTGTGTTGCTTGCACGATG 5400  
5391 ATATGAGAGCAGCCGCTGATTCGAAACCCAGTCCACAGCTATTCTCTGCACCAACTGACC 5450  
5401 ATATGAGAGCAGCCGCTGATTCGAAACCCAGTCCACAGCTATTCTCTGCACCAACTGACC 5460  
5451 TGAAGTTCACTCAGGTCAACCCACAGCTGAGGCGCCAGTGCACACCCCACTGTTT 5510  
5461 TGAAGTTCACTCAGGTCAACCCACAGCTGAGGCGCCAGTGCACACCCCACTGTTT 5520  
5511 AGCTCACTGGAATATCGAGTGGGCTGACCCCAAGGAGAACCGGACCAATGAAGAAA 5570  
5521 AGCTCACTGGAATATCGAGTGGGCTGACCCCAAGGAGAACCGGACCAATGAAGAAA 5580  
5571 TCACACTTCTGAGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5630  
5581 TCACACTTCTGAGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5640  
5631 ATGAGTCACTGAGTCTATGCTCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5690  
5641 ATGAGTCACTGAGTCTATGCTCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5700  
5691 TCACACTTCTGAGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5750  
5701 TCACACTTCTGAGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5760  
5751 CCACATCACTGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5810  
5761 CCACATCACTGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5820  
5811 CCGTTCAGGCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAAGCT 5870  
5821 CCGTTCAGGCAATGGCCAGACTCCAATCCAGAGAACCAATCAAGCCAGATGTCAAGCT 5880  
5871 ACACATCACTGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5930  
5881 ACACATCACTGAGTCACTGCTTAAAGGACACTTTCAGAACAGACAGTCAAGGAGTTG 5940  
5931 ACATGCTCGAGCTCCCTGCTGATCGAGCTCCACCTGCTTCCATGATGATGATGATG 5990  
5941 ACATGCTCGAGCTCCCTGCTGATCGAGCTCCACCTGCTTCCATGATGATGATGATG 6000  
5991 ACCTGCTTCTGAGGACCAACCCAGTCTGCTGATTCATGAGGAGGAGGAGGAGGAGG 6050  
6001 ACCTGCTTCTGAGGACCAACCCAGTCTGCTGATTCATGAGGAGGAGGAGGAGGAGG 6060  
6051 CCAGGATTAACCGGCTACATCAATCAAGTATGAGAGCTGGGCTCTCTCCAGAGAGTGG 6110  
6061 CCAGGATTAACCGGCTACATCAATCAAGTATGAGAGCTGGGCTCTCTCCAGAGAGTGG 6120  
6111 TCCTCGGCGCGCTGCTGATCAGAGGCTTCTTACTTGGCTGGAACCGGAGCCG 6170  
6121 TCCTCGGCGCGCTGCTGATCAGAGGCTTCTTACTTGGCTGGAACCGGAGCCG 6180  
6171 AATATCAATTTATGCTTATGCTTGAAGAAATTAATCAGAGAGGAGGAGGAGGAGGAG 6230  
6181 AATATCAATTTATGCTTATGCTTGAAGAAATTAATCAGAGAGGAGGAGGAGGAGGAG 6240  
6231 GGAAAAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6290





```

; LENGTH: 4027 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 3..4013
; US-08-551-356-5

```

Query Match 28.4%; Score 2211.8; DB 2; Length 4027;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 7; Indels 273;

		Best local similarity	95.99, 1	Identical matches	7; Indels	273; Gaps	1
		Matches	2499; Conservative	0; Mismatches			
QY	3199	CTTTGCACTGAGCCATGGGAGGAGAGCAAGCCTCTGACTGCTCAACAGACAAACCAAACT	3258				
DB	2	CATGGCAGTGAAGTATGGGAGGAGAGCAAGCCTCTGACTGCTCAACAGACAAACCAAACT	61				
QY	3259	GGATGCTCCACATAACCTCCAGTTTGCATGAAGAACTGATTTACTCTCTGCTGGTGAAGT	3318				
DB	62	GGATGCTCCACATAACCTCCAGTTTGCATGAAGAACTGATTTACTCTCTGCTGGTGAAGT	121				
QY	3319	GACTCCACCTCGGSCCCAGATAACAGGATACCGACTGACCGTGGGCTTATCCCGAAGAGG	3378				
DB	122	GACTCCACCTCGGSCCCAGATAACAGGATACCGACTGACCGTGGGCTTATCCCGAAGAGG	181				
QY	3379	ACAGCCAGCAGTACAAATGTGGTCCCTCTGCTTCCAACTACCCACTGAGGAATCTGCA	3438				
DB	182	CCAGCCAGCAGTACAAATGTGGTCCCTCTGCTTCCAACTACCCCTGAGGAATCTGCA	241				
QY	3439	GCCTGCATCTGAGTACACCGTATCCCTCGTGGCCATAAAGGCAACCAAGAGAGCCCCAA	3498				
DB	242	GCCTGCATCTGAGTACACCGTATCCCTCGTGGCCATAAAGGCAACCAAGAGAGCCCCAA	301				
QY	3499	AGCCACTGGAGTCTTTACCACTGCGAGCCTGGAGCTCTATTCCACCTTACAACACCGA	3558				
DB	302	AGCCACTGGAGTCTTTACCACTGCGAGCCTGGAGCTCTATTCCACCTTACAACACCGA	361				
QY	3559	GGTCACTGAGACCAACCAATGTGATCACTGAGACGCCCTGCTCCAAGAAATTGGTTTTAAGCT	3618				
DB	362	GGTCACTGAGACCAACCAATGTGATCACTGAGACGCCCTGCTCCAAGAAATTGGTTTTAAGCT	421				
QY	3619	GGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAACTGATCTTCAGACTCAGGAAG	3678				
DB	422	GGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAACTGATCTTCAGACTCAGGAAG	481				
QY	3679	CATCGTTGTGTCGGCTTGACTCCAGGAGTAGAATAGCTTACACCATCCAAGTCCCTGAG	3738				
DB	482	CATCGTTGTGTCGGCTTGACTCCAGGAGTAGAATAGCTTACACCATCCAAGTCCCTGAG	541				
QY	3739	AGATGGAAGGAAAGAGATGGCCCAATTTGTAACAAGTGGTGTACACCATTTGCTCCACC	3798				
DB	542	AGATGGAAGGAAAGAGATGGCCCAATTTGTAACAAGTGGTGTACACCATTTGCTCCACC	601				
QY	3799	AACAAATTGTCATCTGGAGGCAACCCCTGACACTGGAGTGTCTCAGTCTCCTGGGAG	3858				
DB	602	AACAAATTGTCATCTGGAGGCAACCCCTGACACTGGAGTGTCTCAGTCTCCTGGGAG	661				
QY	3859	GAGCACACCCAGACATTTACTGGTTATAGAAATACCAACACCCCTACAACAGCCGCGACGA	3918				
DB	662	GAGCACACCCAGACATTTACTGGTTATAGAAATACCAACACCCCTACAACAGCCGCGACGA	721				
QY	3919	GGGAAATTTCTTGGAAAGTGGTCCATCTGTGATCAGAGCTCTCTGCACTTTTGTAACT	3978				
DB	722	GGGAAATTTCTTGGAAAGTGGTCCATCTGTGATCAGAGCTCTCTGCACTTTTGTAACT	781				
QY	3979	GAGTCCCGGCTGGAGTACAATGTCAAGTGTTTACACTGTCAAGGATGACAAGGAAGTGT	4038				
DB	782	GAGTCCCGGCTGGAGTACAATGTCAAGTGTTTACACTGTCAAGGATGACAAGGAAGTGT	841				
QY	4039	CCCTATCTGTGATCAACATCAATCCCA	4063				
DB	842	CCCTATCTGTGATCAACATCAATCCCAAGAGGTGCCCACTCACTGACCTTAAGCTTTGTTGA	901				

Qy	4064	-----		4064
Db	902	TATAACCGANTCAAGCATCGGCGCTGAGTGGAGCCCGGTAAACTCTTCACCATTATTTGG	961	
Qy	4064	-----		4063
Db	962	GTACCGCATCACAGTAGTTGCGGCAGAGAAAGTATCCCTATTTTTGAAGATTTTGTA	1021	
Qy	4064	-----		4063
Db	1022	CTCCTCAGTAGGACTACACAGTCA CAGGCTGGAGCGGGCATTGACTATGATATCAG	1081	
Qy	4064	-----		4065
Db	1082	CGTTATCACTCTCATTAATGGGGCGAGAGTGCCTTACTACACTGACACAACAAACGGC	1141	
Qy	4065	TGTTCTCTCTCCCATCGACTCGGATTCCCAAATTTGTTCCAGACACCATGGGTGTCAAC	4125	
Db	1142	TGTTCTCTCTCCCATCGACTCGGATTCCCAAATTTGTTCCAGACACCATGGGTGTCAAC	1201	
Qy	4126	CTGGGCTCAACCCCCATCCATTGATNTTAAACCAACTCTCTGGTGGTTACTCACTCTGTGAA	4185	
Db	1202	CTGGGCTCAACCCCCATCCATTGATNTTAAACCAACTCTCTGGTGGTTACTCACTCTGTGAA	1261	
Qy	4186	AAATGAGGAAGATGTTGCCAGAGTTGTCAATTTCTCTTCAGACAATGCAGTGTCTTAAC	4245	
Db	1262	AAATGAGGAAGATGTTGCCAGAGTTGTCAATTTCTCTTCAGACAATGCAGTGTCTTAAC	1321	
Qy	4246	AAATCTCTGCTGTGTACAGAATPATGTAGTGTCTCCAGTGTCTACGAACAACATGA	4305	
Db	1322	AAATCTCTGCTGTGTACAGAATPATGTAGTGTCTCCAGTGTCTACGAACAACATGA	1381	
Qy	4306	GAGCACACCTCTTTAGAGGAAGACAGAAAACAGAGTCTTGATTCCCCAACTGGCATTTGACTT	4365	
Db	1382	GAGCACACCTCTTTAGAGGAAGACAGAAAACAGAGTCTTGATTCCCCAACTGGCATTTGACTT	1441	
Qy	4366	TTCTGATATTACTGECOAACCTCTTTTACTGTGCACTGTGATTTGCTCTCGAGCCACCATTAC	4425	
Db	1442	TTCTGATATTACTGECOAACCTCTTTTACTGTGCACTGTGATTTGCTCTCGAGCCACCATTAC	1501	
Qy	4426	TGGCTACAGGATCGGCCATCATCCGAGCACTTCAGTGGAGACCTCGAAGAATCGGGT	4485	
Db	1502	TGGCTACAGGATCGGCCATCATCCGAGCACTTCAGTGGAGACCTCGAAGAATCGGGT	1561	
Qy	4486	GCOCCTACTTCGGAAATTCATACCTTCACCAACTCACTCCAGGACACAGATATGTGGT	4545	
Db	1562	GCOCCTACTTCGGAAATTCATACCTTCACCAACTCACTCCAGGACACAGATATGTGGT	1621	
Qy	4546	CAGCATCGTTGTCTTTAATGGCAGAGAGAAAGTCCCCTTATTGATTGGCCCAACAATCAAC	4605	
Db	1622	CAGCATCGTTGTCTTTAATGGCAGAGAGAAAGTCCCCTTATTGATTGGCCCAACAATCAAC	1681	
Qy	4606	AGTTTCGTATGTTCCGAGGACCTGGNAHTTGTGTGCGACCCCAACAGCTACTGAT	4665	
Db	1682	AGTTTCGTATGTTCCGAGGACCTGGNAHTTGTGTGCGACCCCAACAGCTACTGAT	1741	
Qy	4666	CAGCTGGATGCTCCTGCTGTCAAGTGAATATTA CAGGATCACTTACGGAGAAAAACGG	4725	
Db	1742	CAGCTGGATGCTCCTGCTGTCAAGTGAATATTA CAGGATCACTTACGGAGAAAAACGG	1801	
Qy	4726	AGGAAATAGCCCTGTTCAGAGATTCACTGTGCTGGAGCAAGTCTACAGCTACCATCAG	4785	
Db	1802	AGGAAATAGCCCTGTTCAGAGATTCACTGTGCTGGAGCAAGTCTACAGCTACCATCAG	1861	
Qy	4786	CGGCTTTAAACCTGGAGTTGATTATACCATCACTGTGTATGCTCACTGGCCCTGGAGA	4845	
Db	1862	CGGCTTTAAACCTGGAGTTGATTATACCATCACTGTGTATGCTCACTGGCCCTGGAGA	1921	
Qy	4846	CAGCCCGGACGAGCAAGCCAATTTCCATTATTACCGAACAGAAATTCGACCAACCATC	4905	
Db	1922	CAGCCCGGACGAGCAAGCCAATTTCCATTATTACCGAACAGAAATTCGACCAACCATC	1981	
Qy	4906	CCAGATGCAAGTGACCCGATGTTTCAGGACAAACAGCATTAGTGTCAAGTGGCTTCAAG	4955	

Db 1982 CCAGATGCAAGTACCGGATGTTACAGGACACAGCATTTAGTGTCAAGTGGTGGCTTCAAG 2041  
2Y TTCCCTGTTACTGTTACAGAGTAAACACACATCCCAAAAATGGACAGGACCAACAA 5025  
Db 2042 TTCCCTGTTACTGTTACAGAGTAAACACACATCCCAAAAATGGACAGGACCAACAA 2101  
2Y 5026 AACTAAATGTCAGGTCCAGATCAACACAGAAATGACTATTGAAGGCTTGACGCCACAGT 5085  
Db 2102 AACTAAATGTCAGGTCCAGATCAACACAGAAATGACTATTGAAGGCTTGACGCCACAGT 2161  
2Y 5086 GGAGTATGTTAGTGTCTATGCTCAGATCCAGCGGAGAGTCAAGCTTGTGTTCA 5145  
Db 2162 GGAGTATGTTAGTGTCTATGCTCAGATCCAGCGGAGAGTCAAGCTTGTGTTCA 2221  
2Y 5146 GACTGCAATTAACCAATGATGCCCTTAAAGGACTGGCATTCACATGATGTGATGCGA 5205  
Db 2222 GACTGCAATTAACCAATGATGCCCTTAAAGGACTGGCATTCACATGATGTGATGCGA 2281  
2Y 5206 TTCCATCAAAATTCCTTGGGAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTA 5265  
Db 2282 TTCCATCAAAATTCCTTGGGAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTA 2341  
2Y 5266 CTCGAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACCTGATGTTGAAGACAC 5325  
Db 2342 CTCGAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACCTGATGTTGAAGACAC 2401  
2Y 5326 TGCAGACTGCAAGGCTCAGACCGGTTCTGATACACAGTCAAGTGTGTTGCCCTTGA 5385  
Db 2402 TGCAGACTGCAAGGCTCAGACCGGTTCTGATACACAGTCAAGTGTGTTGCCCTTGA 2461  
2Y 5386 CGATGATATGGAGAGCCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCCTGCACCTG 5445  
Db 2462 CGATGATATGGAGAGCCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCCTGCACCTG 2521  
2Y 5446 TGACCTGAAGTTCACTCAGGTCAACCCACAGCCCTGAGCGCCCTGAGTGGACACCAAC 5505  
Db 2522 TGACCTGAAGTTCACTCAGGTCAACCCACAGCCCTGAGCGCCCTGAGTGGACACCAAC 2581  
2Y 5506 TGTTCAGCTCACTGGATATCAGTGGGGTGACCCCAAGGAGAACCCGACCATGAA 5565  
Db 2582 TGTTCAGCTCACTGGATATCAGTGGGGTGACCCCAAGGAGAACCCGACCATGAA 2641  
2Y 5566 AGAAATCAACCTTGTCTCAGAGCTCATCCGTTGTTGATCAGGACTTATGTTGGCCAC 5625  
Db 2642 AGAAATCAACCTTGTCTCAGAGCTCATCCGTTGTTGATCAGGACTTATGTTGGCCAC 2701  
2Y 5626 CAATATGAAGTGAAGTGTCTATGCTTTAAGGACACTTTGAAGCAGACCCAGCTCAGGG 5685  
Db 2702 CAATATGAAGTGAAGTGTCTATGCTTTAAGGACACTTTGAAGCAGACCCAGCTCAGGG 2761  
2Y 5686 AGTTGTCAACCACTCTGGAG 5704  
Db 2762 TGTGTCAACCACTCTGGAG 2780

RESULT 11  
PCT-US93-12687-5  
Sequence 5, Application PC/TUS9312687  
GENERAL INFORMATION:  
APPLICANT: Irani, Meher H.  
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: ZymoGenetics, Inc.  
STREET: 4225 Roosevelt Way, N.E.  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12687  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/998,271  
FILING DATE: 31-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, Gary E  
REGISTRATION NUMBER: 31-648  
REFERENCE/DOCKET NUMBER: 92-28PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-547-8080 ext 322  
TELEFAX: 206-548-2329  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4027 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..4013  
PCT-US93-12687-5

Query Match 28.4%; Score 2211.8; DB 5; Length 4027;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 7; Indels 273; Gaps 1;  
QY 3199 CTTTGCACTGAGCCATGGGAGGAGGAGCAAGCTCTGACTGCTCAACAGACCAACCAACT 3258  
Db 2 CATGCACTGAGTCAATGAGGAGGAGGAGCAAGCTCTGACTGCTCAACAGACCAACCAACT 61  
QY 3259 GGATGCTCCCACTAACTTCACTGTTGTCAATGAATGATTTCTACTGCTCTGTTGAGATG 3318  
Db 62 GGATGCTCCCACTAACTTCACTGTTGTCAATGAATGATTTCTACTGCTCTGTTGAGATG 121  
QY 3319 GACTCCACTCGGGCCGAGATAACAGGATACCGACTGACCGTGGGCTTACCCGAGAGG 3378  
Db 122 GACTCCACTCGGGCCGAGATAACAGGATACCGACTGACCGTGGGCTTACCCGAGAGG 181  
QY 3379 ACAGCCAGGAGTACATGTTGGTCCCTGCTCTCCAAAGTACCCACTGAGGAATCTGCA 3438  
Db 182 CCAGCCAGGAGTACATGTTGGTCCCTGCTCTCCAAAGTACCCCTTGAAGATCTGCA 241  
QY 3439 GCCTGCACTGAGTACACCGTATCCCTGCTGGCCATAAAGGGCAACCAAGAGAGCCCAA 3498  
Db 242 GCCTGCACTGAGTACACCGTATCCCTGCTGGCCATAAAGGGCAACCAAGAGAGCCCAA 301  
QY 3499 AGCCACTGGAGTCTTTACCACTGACGCTGGGAGCTCTATTCCACCTTACACACCGA 3558  
Db 302 AGCCACTGGAGTCTTTACCACTGACGCTGGGAGCTCTATTCCACCTTACACACCGA 361  
QY 3559 GGTGACTGAGACCACTCGTGTATCACAATGAGCGCTGCTCCAAAGATTTGTTTAAAGCT 3618  
Db 362 GGTGACTGAGACCACTCGTGTATCACAATGAGCGCTGCTCCAAAGATTTGTTTAAAGCT 421  
QY 3619 GGTGACTGAGACCACTCGTGTATCACAATGAGCGCTGCTCCAAAGATTTGTTTAAAGCT 3678  
Db 422 GGTGACTGAGACCACTCGTGTATCACAATGAGCGCTGCTCCAAAGATTTGTTTAAAGCT 481  
QY 3679 CATGCTGTTGCTCGGCTTGAATCCAGGAGTAGAATAGCTCTTACACCATCCAAAGTCTGAG 3738  
Db 482 CATGCTGTTGCTCGGCTTGAATCCAGGAGTAGAATAGCTCTTACACCATCCAAAGTCTGAG 541  
QY 3739 AGATGGACAGAAAGAGATGCGCCAAATTTGAACAAAGTGGTGAACACATTTCTCACC 3798  
Db 542 AGATGGACAGAAAGAGATGCGCCAAATTTGAACAAAGTGGTGAACACATTTCTCACC 601  
QY 3799 AACAAATTCATCTGGAGGCAACCCCTGACCTGGAGTGTCTACAGTCTCTCTGGGAGAG 3858  
Db 542 AGATGGACAGAAAGAGATGCGCCAAATTTGAACAAAGTGGTGAACACATTTCTCACC 601

Db	602	AACAACTTGCATCTGGAGGCAACCTGACACTGGAGTGCTCACAGTCTCTCTGGAGAG	661	Qy	4566	CAGCTGGATGCTCTCTGCTGTCAAGTGAATATACAGATCACTTACGGAGAAACAGG	4725
Qy	3859	GAGCACACCCAGACATTAAGTGTATAGAAATACCAACCCCTACAAACGGCCAGCA	3918	Db	1742	CAGCTGGATGCTCTCTGCTGTCAAGTGAATATACAGATCACTTACGGAGAAACAGG	1801
Db	662	GAGCACACCCAGACATTAAGTGTATAGAAATACCAACCCCTACAAACGGCCAGCA	721	Qy	4726	AGGAAATAGCCCTGTCCAGGAGTTCAGTGTGGGAGCAAGTCTACAGCTACCAATCAG	4785
Qy	3919	GGGAAATCTTTGGGAAAGTGGTCCATCTCATCAGAGCTCTGCACTTTTGTAACT	3978	Db	1802	AGGAAATAGCCCTGTCCAGGAGTTCAGTGTGGGAGCAAGTCTACAGCTACCAATCAG	1861
Db	722	GGGAAATCTTTGGGAAAGTGGTCCATCTCATCAGAGCTCTGCACTTTTGTAACT	781	Qy	4786	CGGCTTAAACCTGGAGTGTATATACATCACTGTGTATGTGTCACTGGCCGTGGAGA	4845
Qy	3979	GAGTCCGGCTGGAGTACAATGTAGTGTATACCTGTCAAGAGTACAGAGAAAGTGT	4038	Db	1862	CGGCTTAAACCTGGAGTGTATATACATCACTGTGTATGTGTCACTGGCCGTGGAGA	1921
Db	782	GAGTCCGGCTGGAGTACAATGTAGTGTATACCTGTCAAGAGTACAGAGAAAGTGT	841	Qy	4846	CAGCCCGCAGCAGCAAGCCAAATTTCCATTAATTAACCGAACAGAAATGACAAACCATC	4905
Qy	4039	CCCTATCTCTGATACCATCATCCCA-----	4063	Db	1922	CAGCCCGCAGCAGCAGCAAGCCAAATTTCCATTAATTAACCGAACAGAAATGACAAACCATC	1981
Db	842	CCCTATCTCTGATACCATCATCCCAAGAGTGTCCTCACTGACCTAAGCTTTGTGTA	901	Qy	4906	CCAGATGCAAGTGAACCGATGTTCCAGCAACACAGCATTAAGTGTCAAGTGTGCTTCAAG	4965
Qy	4064	-----	4063	Db	1982	CCAGATGCAAGTGAACCGATGTTCCAGCAACACAGCATTAAGTGTCAAGTGTGCTTCAAG	2041
Db	902	TATAACCGATCAAGCATCGGCTGAGGTGGACCCCGCTAAACTCTTCCACCATTAATGG	961	Qy	4966	TTCCCTGTGTACTGTTCAGAGTACCACTCCCAAAATGACAGCCACCAACAA	5025
Qy	4064	-----	4063	Db	2042	TTCCCTGTGTACTGTTCAGAGTACCACTCCCAAAATGACAGCCACCAACAA	2101
Db	962	GTACCGCATCAGATGTTGGCGAGGAGAGGTATCCCTATTTTGAAGATTTTGTGTA	1021	Qy	5026	AACTAAACTGCAAGTCCAGATCAAAACAGAAATGACTATTGAAGGCTTGACGCCACAGT	5085
Qy	4064	-----	4063	Db	2102	AACTAAACTGCAAGTCCAGATCAAAACAGAAATGACTATTGAAGGCTTGACGCCACAGT	2161
Db	1022	CTCCTCAGTAGGATACACACAGTACAGGGCTGGAGCGGCAATTGACTATGATATCAG	1081	Qy	5086	GGAGTATGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGTCCAGCTCTGTTCA	5145
Qy	4064	-----	4063	Db	2162	GGAGTATGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGTCCAGCTCTGTTCA	2221
Db	1082	CGTATATCTCTCAATTAATGGCGGAGAGTGCCCTCTACTACATGACAAACAAAGGC	1141	Qy	5146	GACTGCAGTAACCAACATTAAGTGTGCTTAAAGGACTTGGCACTTCACTGTGTGATGCA	5205
Qy	4066	TGTTCCCTCCCTCAGCTGACCTCGATTCACCAACATTTGCTCAGACACCATGCTGTAC	4125	Db	2222	GACTGCAGTAACCAACATTAAGTGTGCTTAAAGGACTTGGCACTTCACTGTGTGATGCA	2281
Db	1142	TGTTCCCTCCCTCAGCTGACCTCGATTCACCAACATTTGCTCAGACACCATGCTGTAC	1201	Qy	5206	TTCCATCAAAATTTGCTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTA	5265
Qy	4126	CTGGGCTCACCACCATCAATGATTTAAACAACTTCTGCTGGTGTACTCAGCTGTGAA	4185	Db	2282	TTCCATCAAAATTTGCTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTA	2341
Db	1202	CTGGGCTCACCACCATCAATGATTTAAACAACTTCTGCTGGTGTACTCAGCTGTGAA	1261	Qy	5266	CTCGAGCTTGAGGTGAATCCATGAGCTATTCCCTGACCTGATGTTGAAGAGACAC	5325
Qy	4186	AAATGAGGAAGATGTTGACAGTGTCTCAATTTCTCTTCCAGCAATGAGTGTCTTAAC	4245	Db	2342	CTCGAGCTTGAGGTGAATCCATGAGCTATTCCCTGACCTGATGTTGAAGAGACAC	2401
Db	1262	AAATGAGGAAGATGTTGACAGTGTCTCAATTTCTCTTCCAGCAATGAGTGTCTTAAC	1321	Qy	5326	TGAGAGCTGCAAGGCTCAGACCGGGTTCAGGTACACAGTCAAGTGTGGTTCCTTGCA	5385
Qy	4246	AAATCTCTGCTGTGACAGATATGTAGTGTCTCCAGTGTCTACGAAACAAATGA	4305	Db	2402	TGAGAGCTGCAAGGCTCAGACCGGGTTCAGGTACACAGTCAAGTGTGGTTCCTTGCA	2461
Db	1322	AAATCTCTGCTGTGACAGATATGTAGTGTCTCCAGTGTCTACGAAACAAATGA	1381	Qy	5386	CGATGATATGGAGAGCCAGCCCTGATTTGAAACCCAGTCCACAGCTATTCTCGACCAAC	5445
Qy	4306	GAGCACACCTTTAGAGGAAGACAGAAACAGGTCTTGATTTCCCACTGGCATTTGACTT	4365	Db	2462	CGATGATATGGAGAGCCAGCCCTGATTTGAAACCCAGTCCACAGCTATTCTCGACCAAC	2521
Db	1382	GAGCACACCTTTAGAGGAAGACAGAAACAGGTCTTGATTTCCCACTGGCATTTGACTT	1441	Qy	5446	TGACCTGAAAGTTCATCTCAGGTCAACCCAGAGCTGAGCGCCAGTGAACACCAAC	5505
Qy	4366	TTCTGATATTAAGTCCAACTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCAC	4425	Db	2522	TGACCTGAAAGTTCATCTCAGGTCAACCCAGAGCTGAGCGCCAGTGAACACCAAC	2581
Db	1442	TTCTGATATTAAGTCCAACTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCAC	1501	Qy	5506	TGTTTACCTCAGCTGAGTATCGAGTGGGTGAGCCCGGAGGAGAGCCGAGCCCAATGA	5565
Qy	4426	TGCTTACAGATCCGCTATCCGAGCATTTCACTGGGAGACCTCGAGAGATCGGT	4485	Db	2582	TGTTTACCTCAGCTGAGTATCGAGTGGGTGAGCCCGGAGGAGAGCCGAGCCCAATGA	2641
Db	1502	TGCTTACAGATCCGCTATCCGAGCATTTCACTGGGAGACCTCGAGAGATCGGT	1561	Qy	5566	AGAAATCAACCTTGTCTTGTGACAGCTCATCCGTGGTGTATCAGGACTTATGGTGGCCAC	5625
Qy	4486	GCCCCACTCTCGGAATTCATCACCTCACCACCTCAGTCCAGGACAGAGATGTGGT	4545	Db	2642	AGAAATCAACCTTGTCTTGTGACAGCTCATCCGTGGTGTATCAGGACTTATGGTGGCCAC	2701
Db	1562	GCCCCACTCTCGGAATTCATCACCTCACCACCTCAGTCCAGGACAGAGATGTGGT	1621	Qy	5626	CAAAATGAGTGTGCTTGTCTTAAAGACACTTTGACAGAGCCAGCTCAGG	5685
Qy	4546	CAGCATCTGCTCTTATGGCAGAGAGAAAGTCCCTTATGATTTGCTGACCAATCAAC	4605	Db	2702	CAAAATGAGTGTGCTTGTCTTAAAGACACTTTGACAGAGCCAGCTCAGG	2761
Db	1622	CAGCATCTGCTCTTATGGCAGAGAGAAAGTCCCTTATGATTTGCTGACCAATCAAC	1681	Qy	5686	AGTTGTCAACCTCTGGAG	5704
Qy	4606	AGTTTCTGATTTCCGAGGACCTGGAAGTGTGCTGCGACCCCGCCAGCTTACTGAT	4665	Db	2762	TGTTGTCAACCTCTGGAG	2780
Db	1682	AGTTTCTGATTTCCGAGGACCTGGAAGTGTGCTGCGACCCCGCCAGCTTACTGAT	1741				

## RESULT 12

JS-09-366-009-20

Sequence 20, Application US/09366009

Patent No. 6426042

## GENERAL INFORMATION:

APPLICANT: Asada, Kiyozo

Uemori, Takashi

Koyama, No. 6426042uto

Hashino, Kimikazu

Kato, Ikunoshin

Kato, Ikunoshin

TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET

CELLS WITH RETROVIRUS

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: WEISER &amp; ASSOCIATES

STREET: 230 South Fifteenth Street, Suite 500

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/366,009

FILING DATE: 02-Aug-1999

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/809,156

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: JP 294382/1995

FILING DATE: 13-NOV-1995

APPLICATION NUMBER: JP 051847/1996

FILING DATE: 08-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weiser, Gerard J.

REGISTRATION NUMBER: 19,763

REFERENCE/DOCKET NUMBER: 977.6507P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-875-8383

TELEFAX: 215-875-8394

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 2481 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

JS-09-366-009-20

Query Match

Best Local Similarity 18.0%; Score 1400; DB 4; Length 2481;

Matches 1895; Conservative 0; Mismatches 275; Indels 300; Gaps 3;

4071 CTCTCCCACTGCGATTCCCAACATTGGTCCAGACACCATCGTGTCACTGGG 4130

8 CTAGCCCCACTGCGATTCCCAACATTGGTCCAGACACCATCGTGTCACTGGG 67

4131 CTCACCCCATCCATTGTTAAACCACTTCTGGTGGTGTACTCCTGTGAAATG 4190

68 CTCACCCCATCCATTGTTAAACCACTTCTGGTGGTGTACTCCTGTGAAATG 127

4191 AGGAAGATGTTCAGAGTTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAAACAATC 4250

128 AGGAAGATGTTCAGAGTTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAAACAATC 187

4251 TCCTGGCTGGTACAGAAATATGATGATGTCTCCAGTGTCTACGAAACATGAGGCA 4310

Db 188 TCCTGCCCTGTACAGAAATATGTAGTGTCTCCAGTGTCTACGAAACAAATGAGAGCA 247

QY 4311 CACCTCTTTAGAGGAGACAGAAAACAGGTCTTGATTTCCCAACTGGCATTGACATTTCTG 4370

Db 248 CACCTCTTTAGAGGAGACAGAAAACAGGTCTTGATTTCCCAACTGGCATTGACATTTCTG 307

QY 4371 ATATTACTGCCAACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCCAACCATCTGGCT 4430

Db 308 ATATTACTGCCAACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCCAACCATCTGGCT 367

QY 4431 ACAGGATCCGCGCATCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGGCCC 4490

Db 368 ACAGGATCCGCGCATCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAAGATCGGGTGGCCC 427

QY 4491 ACTCTCGGAATTCATCACTCCCTCAACCACTCTCTCCAGGACAGAGTATGTGTGAGCA 4550

Db 428 ACTCTCGGAATTCATCACTCCCTCAACCACTCTCTCCAGGACAGAGTATGTGTGAGCA 487

QY 4551 TCGTTGCTCTTAATGGCAGAGAGGAAAGTCCCTTATTGATTGGCCAAATCAACAGTTT 4610

Db 488 TCGTTGCTCTTAATGGCAGAGAGGAAAGTCCCTTATTGATTGGCCAAATCAACAGTTT 547

QY 4611 CTGATGTTCCGAGGACCTGGAAAGTTGTTGTTGGACCCCCCAACAGCTTACTGATAGCT 4670

Db 548 CTGATGTTCCGAGGACCTGGAAAGTTGTTGTTGGACCCCCCAACAGCTTACTGATAGCT 607

QY 4671 GGGATGCTCTCTGCTGTACAGTGTGATATTACAGGATCACTTACGGAGAAACAGGAGAA 4730

Db 608 GGGATGCTCTCTGCTGTACAGTGTGATATTACAGGATCACTTACGGAGAAACAGGAGAA 667

QY 4731 ATAGCCCTGTCCAGGAGTTCACTGTGCTGGGACCAAGTCTACAGTACCATCAGCGGCC 4790

Db 668 ATAGCCCTGTCCAGGAGTTCACTGTGCTGGGACCAAGTCTACAGTACCATCAGCGGCC 727

QY 4791 TTAACCTGGAGTTGATTATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGACGCC 4850

Db 728 TTAACCTGGAGTTGATTATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGACGCC 787

QY 4851 CCGAAGCAGCAAGCCCAATTTCCATTAATTAACGAAACAGAAATTGACAAACCAATCC 4906

Db 788 CCGAAGCAGCAAGCCCAATTTCCATTAATTAACGAAACAGAAATTGACAAACCAATCC 847

QY 4907 -----CAGATGCAAGTGACCGATGTTTACGAGCAACAGCAATGATG 4946

Db 848 GCGTATTCTGTGACCAACTGACCTGAAAGTTCACTAGGTCACACCCCAAGCTGAGCG 907

QY 4947 TCAAGTGGCTGCTTCAAGTTCCCTCTTACTGTTTACAGAGTAAACCACTCCCAAAA 5006

Db 908 CCAAGTGGACACCACTCACTGATGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 967

QY 5007 ATGACACGAGACCAACAAACAACTTAAACTGAGTCCAGATCAACAGAAATGACTATTG 5066

Db 968 AGAAGACCGGACCAATGAAAGAAATCAACCTTGTCTCTGACAGCTCATCGTGGTGTAT 1027

QY 5067 AAGCTTTGCAAGCCCAACAGTGGAGTATGTTGTTAGTGTCTATGCTC----- 5111

Db 1028 CAGGACTTATGTTGGTGGCCCAACCAATATGAAAGTGTCTATGCTCTTAAGGACACTTGA 1087

QY 5112 ----- 5111

Db 1088 CAAGCAGACAGCTCAGGGTGTGTTGTCACCTCTGGAGAAATGTGAGCCCAAGAGGG 1147

QY 5112 ----- 5111

Db 1148 CTCGTGTGACAGATGCTACTGTGAGACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGA 1207

QY 5112 ----- 5111

Db 1208 CGATCTGCTTCCAAAGTTGATGCGGTTCAGCAATGCGGAGACTTCCAAATCCAGAGAA 1267

QY 5112 ----- 5111

Db 1268 CCATCAAGCCAGATGTGAGAAAGCTACACCATCAAGGTTTAAACCAAGGCACTGACTACA 1327



QY 5112 -----AGATCCAGCGGAGAGATCGACCTCTGGTTCCAGACTGAG 5153  
 Db 1328 AGATCTACTGTTACACCTTGTAATGACATGTCGAGGCTCCCTGTGGTCTATCGACGCT 1387  
 QY 5154 TAACCAACATTGATCGCCCTAAAGAGACTGGCATTCACCTGATGTGGATGTGATTCATCA 5213  
 Db 1388 CCACCTGCCATTGATCGACCATCCACCTCGGTTTCTCGGCCACACACCCCAATTCCTTGC 1447  
 QY 5214 AAATTGCTTGGGAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTACCTACTCGAGCC 5273  
 Db 1448 TGGTATATGCGAGCGGCCGATGTCGCGGATATACCGGCTATCATCATCAAGTATGAGAAGC 1507  
 QY 5274 CTGAGGATGGAATCATGAGCTATTCCCTGCACTGATGTTGGAAGAGACACACTCGACAGC 5333  
 Db 1508 CTGGGTCTCTCCACAGAGAGTGGTCTCTCGGCCCGCCCTGGTGTACAGAGGCTACTA 1567  
 QY 5334 TCCAGAGCTCAGACCGGGTCTGAGTACACAGTCAGTGTGGTGTGGTGTGACAGATATA 5393  
 Db 1568 TTACTGGGCTGGAACCGGGAACCGGAATATACAATTTATGTCAITGGCCCTGAAGAAATATC 1627  
 QY 5394 TGGAGAGCGAGCCCTGATTGGAACCC---AGTCCACAGCTATTCTGACCAACTGACC 5450  
 Db 1628 AGAAGAGCGAGCCCTGATTGGAAGGAAGAGCTAGCGCTATTCTGACCAACTGACC 1687  
 QY 5451 TGAAGTTCACTGAGTCAACCCACAGCCTGAGCGCCAGTGGACACACCCCAATGTTTC 5510  
 Db 1688 TGAAGTTCACTGAGTCAACCCACAGCCTGAGCGCCAGTGGACACACCCCAATGTTTC 1747  
 QY 5511 AGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAA 5570  
 Db 1748 AGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAA 1807  
 QY 5571 TCAACCTTGCTCTGACAGCTCATCCGTTGTTATCAGGACTTATGGTGGCCACCAAAAT 5630  
 Db 1808 TCAACCTTGCTCTGACAGCTCATCCGTTGTTATCAGGACTTATGGTGGCCACCAAAAT 1867  
 QY 5631 ATGAAGTGTGTTCTATGCTCTTAAGGACACTTTGACAGAGACAGCCAGCTCAGGAGTTG 5690  
 Db 1868 ATGAAGTGTGTTCTATGCTCTTAAGGACACTTTGACAGAGACAGCCAGCTCAGGAGTTG 1927  
 QY 5691 TCACCACTCTGAGAAATGTACGCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGA 5750  
 Db 1928 TCACCACTCTGAGAAATGTACGCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGA 1987  
 QY 5751 CCACCATCACCATTAGCTGGAGAACCAAGACTGAGACCATCACTGGCTCCAGTTGATG 5810  
 Db 1988 CCACCATCACCATTAGCTGGAGAACCAAGACTGAGACCATCACTGGCTCCAGTTGATG 2047  
 QY 5811 CGGTTCCAGCCAAATGGCCAGACTCCCAATCCAGAGAACCATCAAGCCAGATGTCAGAAAGCT 5870  
 Db 2048 CGGTTCCAGCCAAATGGCCAGACTCCCAATCCAGAGAACCATCAAGCCAGATGTCAGAAAGCT 2107  
 QY 5871 ACACCATCACTGGCTTACACAGGACCTGACTACAGATCTACTGTACACCTTGAATG 5930  
 Db 2108 ACACCATCACTGGCTTACACAGGACCTGACTACAGATCTACTGTACACCTTGAATG 2167  
 QY 5931 ACAATGCTGGAGCTCCCTGTGGTTCATCGAGCGCTCCACCTGCCATTTGATGACCAATCCA 5990  
 Db 2168 ACAATGCTGGAGCTCCCTGTGGTTCATCGAGCGCTCCACCTGCCATTTGATGACCAATCCA 2227  
 QY 5991 ACTGGTTTCTGGCCACCAACCCAAATTCCTTGTGTTATCATGGAGCGCCGACGCTG 6050  
 Db 2228 ACTGGTTTCTGGCCACCAACCCAAATTCCTTGTGTTATCATGGAGCGCCGACGCTG 2287  
 QY 6051 CCAGATTAACCGGCTACATCATCAAGTATGAGAGCTTGGGTCTCTCCCGAGAGAGTGG 6110  
 Db 2288 CCAGATTAACCGGCTACATCATCAAGTATGAGAGCTTGGGTCTCTCCCGAGAGAGTGG 2347  
 QY 6111 TCCCTGGGCCCGCCCTGGTGTCAAGAGGCTACTATTACTGGGCTGGAACCGGGAACCG 6170  
 Db 2348 TCCCTGGGCCCGCCCTGGTGTCAAGAGGCTACTATTACTGGGCTGGAACCGGGAACCG 2407

QY 6171 AATATCAATTTATGTCATTGCGCTGAGAGATTAATCAGAGAGCGAGCCCTGATTGGAA 6230  
 Db 2408 AATATCAATTTATGTCATTGCGCTGAGAGATTAATCAGAGAGCGAGCCCTGATTGGAA 2467  
 QY 6231 GGAAAAAGAC 6240  
 Db 2468 GGAAAAAGAC 2477

RESULT 13  
 US-08-809-156B-20  
 ; Sequence 20, Application US/08809156B  
 ; Patent No. 6472204  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Asada, Kiyoze  
 ; APPLICANT: Uemori, Takashi  
 ; APPLICANT: Ueno, Takashi  
 ; APPLICANT: Koyama, No. 6472204auto  
 ; APPLICANT: Hashino, Kimikazu  
 ; APPLICANT: Kato, Ikunoshin  
 ; TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET  
 ; TITLE OF INVENTION: CELLS WITH RETROVIRUS  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WEISER & ASSOCIATES  
 ; STREET: 230 South Fifteenth Street, Suite 500  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19102  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/809,156B  
 ; FILING DATE: 07-MAR-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/JP96/03254  
 ; FILING DATE: 07-NOV-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 294382/1995  
 ; FILING DATE: 13-NOV-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 051847/1996  
 ; FILING DATE: 08-MAR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Weiser, Gerard J.  
 ; REGISTRATION NUMBER: 19,763  
 ; REFERENCE/DOCKET NUMBER: 977,6507P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-875-8383  
 ; TELEFAX: 215-875-8394  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2481 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; US-08-809-156B-20

Query Match 18.0%; Score 1400; DB 4; Length 2481;  
 Best Local Similarity 76.7%; Pred. No. 0;  
 Matches 1895; Conservative 0; Mismatches 275; Indels 300; Gaps 3;  
 QY 4071 CTCCTCCGACTGACCTGCGATTCCACCAATGTTGTCACAGACACCATGCGTGTCACTGGG 4130  
 Db 8 CTAGCCCCACTGCGATTCCACCAATGTTGTCACAGACACCATGCGTGTCACTGGG 67  
 QY 4131 CTCACCCCCCTCATGTTATTTAAACCACTTCTCGTGGTTCCTCACTGCTGTGAAAAATG 4190





1317 GTACACCTTGAATGACATGCTCGGAGCTCCCTGGTGTATCATCGAGCGCTCCACTGCCAT 1376  
5977 TGATGACCATCCAACTGGTTCCTGGSCACACACCAATTCCTGCTGTATCATG 6036  
1377 TGATGACCATCCAACTGGTTCCTGGSCACACACCAATTCCTGCTGTATCATG 1436  
6037 GCAGCCGACGCTGCGAGGATACCGGTACATCATCAAGTATGAGAGCCTGGGTCTCC 6096  
1437 GCAGCCGACGCTGCGAGGATACCGGTACATCATCAAGTATGAGAGCCTGGGTCTCC 1496  
6097 TCCAGAGAGTGGTCCCTCGGCGCCGCTGTGTACAGAGGCTACTATTACTGGCT 6156  
1497 TCCAGAGAGTGGTCCCTCGGCGCCGCTGTGTACAGAGGCTACTATTACTGGCT 1556  
6157 GGAACCGGACCGAATATCAATTTATGCTATGCTGCTGCTGAGAGGCTGAGAGGGA 6216  
1557 GGAACCGGACCGAATATCAATTTATGCTATGCTGCTGCTGAGAGGCTGAGAGGGA 1616  
6217 GCGCTGATTTGGAAGGAAAGAAC---AGACGAGCTTCCCAACTGGTAACCTTCCACA 6273  
1617 GCGCTGATTTGGAAGGAAAGAAC---AGACGAGCTTCCCAACTGGTAACCTTCCACA 1676  
6274 CCCAATCTTCATGACGACGAGATCTTGGATGTTCTTCCAC 6315  
1677 CCCAATCTTCATGACGACGAGATCTTGGATGTTCTTCCAC 1718

## RESULT 15

US-08-809-156B-33

Sequence 33, Application US/08809156B

Patent No. 6472204

## GENERAL INFORMATION:

APPLICANT: Asada, Kiyozo

APPLICANT: Uemori, Takashi

APPLICANT: Ueno, Takashi

APPLICANT: Koyama, No. 6472204uto

APPLICANT: Hashino, Kimikazu

APPLICANT: Kato, Ikunoshin

TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET

TITLE OF INVENTION: CELLS WITH RETROVIRUS

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: WEISER &amp; ASSOCIATES

STREET: 230 South Fifteenth Street, Suite 500

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/809,156B

FILING DATE: 07-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03254

FILING DATE: 07-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 294382/1995

FILING DATE: 13-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 051847/1996

FILING DATE: 08-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weiser, Gerard J.

REGISTRATION NUMBER: 19,763

REFERENCE/DOCKET NUMBER: 977.6507P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-875-8383

TELEFAX: 215-875-8394

## INFORMATION FOR SEQ ID NO: 33:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1722 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-809-156B-33

Query Match 11.8%; Score 921.2; DB 4; Length 1722;

Best Local Similarity 79.5%; Pred. No. 2.1e-253;

Matches 1130; Conservative 0; Mismatches 283; Indels 9; Gaps 3;

QY 4900 ACCATCCAGATGCAAGTGCAGGATGTTCCAGGACACAGCATAGTGTCAAGTGGCTGCC 4959  
DB 300 ACCAAGAGGGCTCGTGTGACGATGCTACTGAGACCAACCATCACATTTAGCTGGAGAAC 359  
QY 4960 TTCAAGTTCCCTCTTACTGTTACAGAGTAACCAACCTCCCAAAAATGGACAGGACC 5019  
DB 360 CAAGACTGAGACGATCACTGGCTTCCAAGTTGATGCGTTCCAGCCAATGGCCAGACTCC 419  
QY 5020 AACAAAACCTAAACTGCAGGTCCAGATCAACAGAAATGACTATTGAAGCTTGCAGCC 5079  
DB 420 AATCCA---GAGAACCATCAAGCCAGATGTCAGAAAGCTACACCATCAGGTTTACAAAC 476  
QY 5080 CACAGTGGAGTATGTTGTTAGTGTCTATGCTTCAAGATCCAAAGCGGAGAGAGTCAAGCTCT 5139  
DB 477 AGGCACCTGACTACAAGATCTACCTGTACACTTGAATGACAAATGCTCGAGCTCCCTGT 536  
QY 5140 GGTTCAGACTGCAGTAACCAACATGATCGCCCTAAAGGACTGGCATTCATGATGTGA 5199  
DB 537 GGTATCGACGCCCTCCATGCCATTTGATGACCACTCCAACTCGGTTTCTTGGCCACCAAC 596  
QY 5200 TGTGATTTCATCAAAATTTGCTTGGGAAAGCCCAAGGCGGCAAGTTTCCAGGTACAGGGT 5259  
DB 597 ACCCAATTCCTTGTGTATCATGSCAGCGCCACGCTGCCAGGATTAACGGCTACATCAT 656  
QY 5260 GACCTACTCGAGCCCTCAGAGTGAATTCATGAGCTATTCCTGACCTGATGGTGAAGA 5319  
DB 657 CAAGTATGAGAAGCCTGGTCTCTCCAGAGAAAGTGTCTCCCTGGCCCGCCCTGGTGT 716  
QY 5320 AGACACTGCAGAGCTGCAAGGCTCAGACCGGGTTCGAGTACACAGTCAAGTGTGGTTGC 5379  
DB 717 CACAGAGCTACTATTACTGGCTTGGAAACCGGNAACCGAATATACATTTATGTCATTGC 776  
QY 5380 CTTGCACGATGATATGAGAGCCAGCCCTGATTTGGAACCC---AGTCCACAGCTATTCC 5436  
DB 777 CTTGAAGAATAATCAGAAAGAGCGGCCCTGATTGGAAGGAAAGAAAGACTAGCGCTATTCC 836  
QY 5437 TGCACCAACTGACCTGAAGTTCACTCAGGTACACCCACCAAGCCTGAGCGCCAGTGGAC 5496  
DB 837 TGCACCAACTGACCTGAAGTTCACTCAGGTACACCCACCAAGCCTGAGCGCCAGTGGAC 896  
QY 5497 ACCACCCAATGTTTCAGCTCACTGATATCGAGTGGGGTGACCCCAAGGAGAGAACCGG 5556  
DB 897 ACCACCCAATGTTTCAGCTCACTGATATCGAGTGGGGTGACCCCAAGGAGAGAACCGG 956  
QY 5557 ACCAATGAAGAAATCAACCTTGTCTCTGACAGCTCATCGTGGTTGTATCAGGACTTAT 5616  
DB 957 ACCAATGAAGAAATCAACCTTGTCTCTGACAGCTCATCGTGGTTGTATCAGGACTTAT 1016  
QY 5617 GGTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAAAGGACACTTTTGACAGAGACC 5676  
DB 1017 GGTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAAAGGACACTTTTGACAGAGACC 1076  
QY 5677 AGCTCAGGGAGTTTGTACCACTCTGGGAATGTGAGCCCAACCAAGAGGGCTGGTGGAC 5736  
DB 1077 AGCTCAGGGTGTGTCTACCACTCTGGGAATGTGAGCCCAACCAAGAGGGCTGGTGGAC 1136  
QY 5737 AGATGCTACTGAGACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGG 5796  
DB 1137 AGATGCTACTGAGACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGG 1196

QY	597	CCTCCAGTTGATGCGTTCCAGCCATGCGCAGACTCCCAATCCAGAGACCCTCAAGCC	5856
Db	1197	CTTCCAAGTTGATGCGTTCCAGCCATGCGCAGACTCCCAATCCAGAGACCCTCAAGCC	1256
QY	5857	AGATGTCAGAGCTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACCT	5916
Db	1257	AGATGTCAGAGCTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACCT	1316
QY	5917	GTACACCTTGAATGACATGCTCGGAGCTCCCTGTGGTATCGAGCCTCCACTGGCAT	5976
Db	1317	GTACACCTTGAATGACATGCTCGGAGCTCCCTGTGGTATCGAGCCTCCACTGGCAT	1376
QY	5977	TGATGCACCATCCAAACCTGCGTTTCTGSCCACCACACCCCAATTCCTTGTGGTATCATG	6036
Db	1377	TGATGCACCATCCAAACCTGCGTTTCTGSCCACCACACCCCAATTCCTTGTGGTATCATG	1436
QY	6037	GCAGCGGCACGTGCCAGATTACCGGCTACATCATCAAGTATGAGAGCCTGGTCTCC	6096
Db	1437	GCAGCGGCACGTGCCAGATTACCGGCTACATCATCAAGTATGAGAGCCTGGTCTCC	1496
QY	6097	TCCCAGAGAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTATTACTGGCCT	6156
Db	1497	TCCCAGAGAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTATTACTGGCCT	1556
QY	6157	GGAAACGGGAACCGAATATACAAATTATGTCTATTCCCTGAGAAATATCAGAGAGCGA	6216
Db	1557	GGAAACGGGAACCGAATATACAAATTATGTCTATTCCCTGAGAAATATCAGAGAGCGA	1616
QY	6217	GCCCTGTATTGGAAGGAAAAAGAC--AGACGAGCTTCCCAACTGGTAACCTTCCACA	6273
Db	1617	GCCCTGTATTGGAAGGAAAAAGAC--AGACGAGCTTCCCAACTGGTAACCTTCCACA	1676
QY	6274	CCCCAATCTTCATGACCCAGAGATCTTGGATGTTCCCTTCCAC	6315
Db	1677	CCCCAATCTTCATGACCCAGAGATCTTGGATGTTCCCTTCCAC	1718

Search completed: March 7, 2004, 21:48:28  
Job time : 595 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: March 7, 2004, 05:41:46 ; Search time 2538 Seconds

(without alignments)  
13047.559 Million cell updates/sec

file: US-10-084-817-2

effect score: 7795

sequence: 1 cgcctcgagccgctgtgct.....ttttaaatatccacagtact 7795

scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

searched: 3373863 seqs, 2124093041 residues

total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Match	Length	DB	ID	Description
1	7780.6	99.8	8044	6	AAS94866	Human DNA
2	7774.8	99.7	8137	4	AAS94866	Human pol
3	7671.8	98.4	8230	4	AAS94866	Human pol
4	7614.8	97.7	8062	4	AAS94866	Human pol
5	7423	95.2	7679	9	ADB31322	Testoster
6	7423	95.2	7680	3	AAP21131	Human low
7	7423	95.2	7680	5	ABR2689	Fibronect
8	7423	95.2	7680	6	ABL7540	Thyroid c
9	7423	95.2	7680	6	ABT11082	Human bre
10	7423	95.2	7680	7	ABZ96825	Human nuc
11	7423	95.2	7680	7	ACF03878	Human fib
12	7423	95.2	7680	7	ABX10391	DNA encod
13	7423	95.2	7680	7	ACC46009	Human fib
14	7423	95.2	7680	7	ACF12859	Human car
15	7423	95.2	7680	7	ACA44817	Human fib
16	7423	95.2	7680	7	ACA4819	Human fib
17	7423	95.2	7680	8	ADB70377	Fibronect
18	7423	95.2	7680	9	ADB98703	Human fib
19	7423	95.2	7680	9	ADD18771	Human dis
20	7423	95.2	7680	9	ADB2499	Human DNA
21	7423	95.2	14740	3	AAS35010	Human ade
22	7423	95.2	14740	3	AAF21132	Human low
23	7423	95.2	14740	7	ABZ96826	Human nuc

24	7419.8	95.2	7680	2	AAT17551	Human fib
25	7418.2	95.2	7705	1	AAQ70596	Sequence
26	7415	95.1	7705	1	AAQ15214	Human fib
27	7410.4	95.1	8066	4	AAI59591	Human pol
28	7410.4	95.1	8066	4	AAI59589	Human pol
29	7410.4	95.1	8066	4	AAI59588	Human pol
30	7410.4	95.1	8066	4	AAI59590	Human pol
31	7341.6	94.2	7705	2	AAQ05502	Sequence
32	7237	92.8	8039	3	AAC98889	Human FN
33	7234.8	92.7	7867	4	AAI57803	Human pol
34	7223	92.7	7803	2	AAQ70009	Fibrinoge
35	7218	92.6	7550	7	ACC00412	Human cel
36	6964	89.3	6988	8	ACD06169	Human CDN
37	6844.4	87.8	8216	5	AAS86466	DNA encod
38	5869.4	75.3	7361	8	ACD06170	Human CDN
39	5320.6	68.3	8329	6	ABK63824	Rat seque
40	5320.6	68.3	8329	9	ADB58331	Toxicity-
41	5320.6	68.3	8329	9	ADB52891	Primary r
42	5185.2	66.5	6816	7	ABX74443	Human CDN
43	3466.2	44.5	4860	3	AAA35009	Human ade
44	2963.8	38.0	3522	5	AAS86462	DNA encod
45	2211.8	28.4	4027	2	AAQ70007	Tissue-bi

## ALIGNMENTS

### RESULT 1

AAS94866

ID AAS94866 standard; DNA; 8044 BP.

XX AC AAS94866;

XX DT 14-FEB-2002 (first entry)

XX DE Human DNA sequence #121 expressed during foam cell differentiation.

XX KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;

XX KW cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200177389-A2.

XX PD 18-OCT-2001.

XX PF 04-APR-2001; 2001WO-US0111128.

XX PR 05-APR-2000; 2000US-0195106P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;

XX PT Tai J;

XX DR WPI; 2002-010925/01.

XX PT Composition useful for diagnosis of conditions, disorders or diseases

XX PT associated with atherosclerosis, comprises several polynucleotides that

XX PT are differentially expressed in foam cell development.

XX PS Claim 1; Page 175-177; 315pp; English.

XX CC The present invention relates to the isolation of human polynucleotide

XX CC sequences that are differentially expressed during foam cell

XX CC differentiation. The polynucleotide sequences of the invention or a

XX CC composition comprising these polynucleotides are useful as a high

XX CC throughput method for detecting altered expression of one or more

XX CC polynucleotides in a sample. The polynucleotides can be used in the

XX CC diagnosis of disorders associated with foam cell development such as

XX CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as

XX CC coronary artery disease. The polynucleotide sequences can also be used as

XX CC PCR primers and probes. The polynucleotides of the invention are also

CC useful in gene therapy. AAS94746-AAS95021 represent the human  
CC polynucleotide sequences of the invention which are differentially  
CC expressed during foam cell differentiation

Sequence 8044 BP: 2143 A: 2107 C: 1977 G: 1817 T: 0 U: 0 Other: 0



Db 4201 TGCAGAGTTGTCAATTTCTCCTTCAGACAATGACAGTGGTCTTAACAAATCTCCTCGCCTGG 4260  
Qy 4261 TACAGAAATATGTAGTAGTGTCTCCAGTGTCTACGAAACAATAGAGACACACCTCTTTAG 4320  
Db 4262 TACAGAAATATGTAGTAGTGTCTCCAGTGTCTACGAAACAATAGAGACACACCTCTTTAG 4320  
Qy 4321 AGCAAGACAGAAAAACAGAGTCTTTGATTCCCACTGGCAATTTCTGATATTTACTGC 4380  
Db 4322 AGCAAGACAGAAAAACAGAGTCTTTGATTCCCACTGGCAATTTCTGATATTTACTGC 4380  
Qy 4381 CAACTCTTTTACTGTGCACTGGATTGCTCTTGAGGCCACCACTACTGGCTACAGGATCCG 4440  
Db 4382 CAACTCTTTTACTGTGCACTGGATTGCTCTTGAGGCCACCACTACTGGCTACAGGATCCG 4440  
Qy 4441 CCAATCATCCGAGCACTTCAGTGGAGACCTCGAAGATCGGGTCCCACTCTCGGAA 4500  
Db 4442 CCAATCATCCGAGCACTTCAGTGGAGACCTCGAAGATCGGGTCCCACTCTCGGAA 4500  
Qy 4501 TTCCATCACTCCACCACTCACTCCAGGCACAGATATGTGTGACATCGTTGCTCT 4560  
Db 4502 TTCCATCACTCCACCACTCACTCCAGGCACAGATATGTGTGACATCGTTGCTCT 4560  
Qy 4561 TAAATGCGAGAGGAAAGTCCCTTATTGATTGGCCAAACAATCAAGTTTCTGATGTTC 4620  
Db 4562 TAAATGCGAGAGGAAAGTCCCTTATTGATTGGCCAAACAATCAAGTTTCTGATGTTC 4620  
Qy 4621 GAGGACCTGGAAGTTGTTGCTGCGACCCCACTCACTGATCGGATGTCTCC 4680  
Db 4622 GAGGACCTGGAAGTTGTTGCTGCGACCCCACTCACTGATCGGATGTCTCC 4680  
Qy 4681 TGCTGTCACTGAGATATTACAGATCACTTACGAGAAACAGGAGAAATAGCCCTGT 4740  
Db 4682 TGCTGTCACTGAGATATTACAGATCACTTACGAGAAACAGGAGAAATAGCCCTGT 4740  
Qy 4741 CCAGGAGTTCACTGTGCTGGAGCAAGTCTACAGTACCATCAGCGGCTTAAACCTGG 4800  
Db 4742 CCAGGAGTTCACTGTGCTGGAGCAAGTCTACAGTACCATCAGCGGCTTAAACCTGG 4800  
Qy 4801 AGTTGATTATACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAG 4860  
Db 4802 AGTTGATTATACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAG 4860  
Qy 4861 CAAAGCAATTTCCATTAATACCGAACAAGAAATGACAAACATCCCAAGATGCAAGTGAC 4920  
Db 4862 CAAAGCAATTTCCATTAATACCGAACAAGAAATGACAAACATCCCAAGATGCAAGTGAC 4920  
Qy 4921 CGATGTTACGACAAACAGCATTTAGTGTCAAGTGGTGGCTTCAAGTTCCCTGTTACTGG 4980  
Db 4922 CGATGTTACGACAAACAGCATTTAGTGTCAAGTGGTGGCTTCAAGTTCCCTGTTACTGG 4980  
Qy 4981 TTACAGAGTTACCACTCTCCCAAAATGGACCGACCAACAAACTAAACTGACAGG 5040  
Db 4982 TTACAGAGTTACCACTCTCCCAAAATGGACCGACCAACAAACTAAACTGACAGG 5040  
Qy 5041 TCCAGATCAAAACAGAAATGACTATTGAAGGCTTGAGCCCACTGAGTATGTGGTTAG 5100  
Db 5042 TCCAGATCAAAACAGAAATGACTATTGAAGGCTTGAGCCCACTGAGTATGTGGTTAG 5100  
Qy 5101 TGCTTATGCTCAGATCCAGCGGAGAGTCAAGCTCTGTTGAGTCAAGTATGTTGGTTAG 5160  
Db 5102 TGCTTATGCTCAGATCCAGCGGAGAGTCAAGCTCTGTTGAGTCAAGTATGTTGGTTAG 5160  
Qy 5161 CATTGATCGCCCTAAAGGACTGGCAATTCAGTGTGATGTGATTCATCCATAAATTCG 5220  
Db 5162 CATTGATCGCCCTAAAGGACTGGCAATTCAGTGTGATGTGATTCATCCATAAATTCG 5220  
Qy 5221 TTGGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTTGACCTTCTGAGCCCTGAGGA 5280  
Db 5222 TTGGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTTGACCTTCTGAGCCCTGAGGA 5280  
Qy 5281 TGGAAATCCATGAGCTATTCTCCCTGACCTGTAGTGGTGAAGAAGACACTGACAGGTCGAGG 5340  
Db 5282 TGGAAATCCATGAGCTATTCTCCCTGACCTGTAGTGGTGAAGAAGACACTGACAGGTCGAGG 5340

Qy 5341 CCTCAGACCGGGTCTGAGTACACAGTCAAGTGTGTTGCTTGCACGATGATGAGAG 5400  
Db 5342 CCTCAGACCGGGTCTGAGTACACAGTCAAGTGTGTTGCTTGCACGATGATGAGAG 5400  
Qy 5401 CCAGCCCTGATTGGAAACCCAGTCCACAGCTATTCTGTGACCAACTGACCTGAAGTTTAC 5460  
Db 5402 CCAGCCCTGATTGGAAACCCAGTCCACAGCTATTCTGTGACCAACTGACCTGAAGTTTAC 5460  
Qy 5461 TCAGGTTCACACCCCAAGCCTGAGCGCCAGTGGACACCAACCAATGTTTCACTGCTACTGG 5520  
Db 5462 TCAGGTTCACACCCCAAGCCTGAGCGCCAGTGGACACCAACCAATGTTTCACTGCTACTGG 5520  
Qy 5521 ATATCAGTGGGGTGCACCCCAAGGAGAGACCGGACCAATGAAGAATCAACCTTCC 5580  
Db 5522 ATATCAGTGGGGTGCACCCCAAGGAGAGACCGGACCAATGAAGAATCAACCTTCC 5580  
Qy 5581 TCTGTACAGCTCATCGTGGTGTATCAGGACTTATGTTGGGCCACCAATATGAAGTGAG 5640  
Db 5582 TCTGTACAGCTCATCGTGGTGTATCAGGACTTATGTTGGGCCACCAATATGAAGTGAG 5640  
Qy 5641 TGCTATGCTCTTAAGGACACTTTGACAGACAGACAGCTCAGGGAGTTGTCCACACTCT 5700  
Db 5642 TGCTATGCTCTTAAGGACACTTTGACAGACAGACAGCTCAGGGAGTTGTCCACACTCT 5700  
Qy 5701 GGAGAAATGTGAGCCCAAGAGGAGGCTGTGTGACAGATGCTACTGAGACCAACCATCAC 5760  
Db 5702 GGAGAAATGTGAGCCCAAGAGGAGGCTGTGTGACAGATGCTACTGAGACCAACCATCAC 5760  
Qy 5761 CATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGGTTCCAGC 5820  
Db 5762 CATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGGTTCCAGC 5820  
Qy 5821 CAAATGCGCAGATCCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAAAGTACACCATCAC 5880  
Db 5822 CAAATGCGCAGATCCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAAAGTACACCATCAC 5880  
Qy 5881 TGCTTTACCAACAGGACAGTACTAAGATCTACCTGTGACACTTGAATGACAAATGCTCG 5940  
Db 5882 AGGTTTACCAACAGGACAGTACTAAGATCTACCTGTGACACTTGAATGACAAATGCTCG 5940  
Qy 5941 GAGTCCCTGTGGTTCATCGAGCGCTTCCACTGCTGATGATGATGATGATGATGATGATGAT 6000  
Db 5942 GAGTCCCTGTGGTTCATCGAGCGCTTCCACTGCTGATGATGATGATGATGATGATGATGAT 6000  
Qy 6001 CTTGGCCACACACCCCAATTCCTTGTGTATCATGGCCCTCACTGCTGATGATGATGATGATGAT 6060  
Db 6002 CTTGGCCACACACCCCAATTCCTTGTGTATCATGGCCCTCACTGCTGATGATGATGATGATGATGAT 6060  
Qy 6061 CGGCTACATCATCAAGTATGAGAGGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGC 6120  
Db 6062 CGGCTACATCATCAAGTATGAGAGGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGC 6120  
Qy 6121 CGGCTGTGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAACCGGAATATACAAAT 6180  
Db 6122 CGGCTGTGTGTACAGAGGCTACTATTACTGGCTGGAAACCGGGAACCGGAATATACAAAT 6180  
Qy 6181 TTATGTCATTGCTCCCTGAAGAAATTAATCAGAGAGCGAGCCCTGATTTGGAAGGAAAGAG 6240  
Db 6182 TTATGTCATTGCTCCCTGAAGAAATTAATCAGAGAGCGAGCCCTGATTTGGAAGGAAAGAG 6240  
Qy 6241 AGACGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCTGAGACGAGATCTT 6300  
Db 6242 AGACGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCTGAGACGAGATCTT 6300  
Qy 6301 GGATGTCCTTCCAGTTCAAAAGACCCCTTGTCAACCCCTTGTGATGAGTATGACACTGG 6360  
Db 6302 GGATGTCCTTCCAGTTCAAAAGACCCCTTGTGATGAGTATGACACTGG 6360  
Qy 6361 AAATGTTATTCAGTCTTCCCTGAGCACTTCTGTTGAGCAACCCAGTGTGAGCAACCAATGAT 6420  
Db 6362 AAATGTTATTCAGTCTTCCCTGAGCACTTCTGTTGAGCAACCCAGTGTGAGCAACCAATGAT 6420











Db 4184 ATTCAACCAATTGGTTCAGACACCATGGGTGTACCTGGGCTCCACCCCATCCATTGA 4243  
Qy 4150 TTATAACCAACTCTCCGTGCGTTACTCACTGTGAAAAATGAGAAAGATGTTGAGAGTT 4209  
Db 4244 TTATAACCAACTCTCCGTGCGTTACTCACTGTGAAAAATGAGAAAGATGTTGAGAGTT 4303  
Qy 4210 GTCAATTTCTCCTTCAGACAATGCAAGTGTCTTAAACAATCTCCTGCTGGTACAGATA 4269  
Db 4304 GTCAATTTCTCCTTCAGACAATGCAAGTGTCTTAAACAATCTCCTGCTGGTACAGATA 4363  
Qy 4270 TGTAGTGTGTCTCCAGTGTCTACGAAACAATGAGAGCACTCTTTAGAGGAACA 4329  
Db 4364 TGTAGTGTGTCTCCAGTGTCTACGAAACAATGAGAGCACTCTTTAGAGGAACA 4423  
Qy 4330 GAAACAGGTCTTGAATCCCAACTGGGATGACATTTCTGATATTACTGCGCACTCTTT 4389  
Db 4424 GAAACAGGTCTTGAATCCCAACTGGGATGACATTTCTGATATTACTGCGCACTCTTT 4483  
Qy 4390 TACTGTGCACTGGATTGCTCTCGAGCACCATCACTGCTACAGGATCCGCCATCATCC 4449  
Db 4484 TACTGTGCACTGGATTGCTCTCGAGCCACCATCACTGCTACAGGATCCGCCATCATCC 4543  
Qy 4450 CGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCGCACTCTCGGAATTCATCAC 4509  
Db 4544 CGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCGCACTCTCGGAATTCATCAC 4603  
Qy 4510 CCTCACCACTCTCACTCCAGGACAGAGATGCTGCTAGCATCGTTGCTTTAATGGCAG 4569  
Db 4604 CCTCACCACTCTCACTCCAGGACAGAGATGCTGCTAGCATCGTTGCTTTAATGGCAG 4663  
Qy 4570 AGAGAAAGTCCCTTATTGATGGCAACAATCAACAGATTTCTGATGTTCCAGGACCT 4629  
Db 4664 AGAGAAAGTCCCTTATTGATGGCAACAATCAACAGATTTCTGATGTTCCAGGACCT 4723  
Qy 4630 GGAAGTGTGCTGCGACCCCAACAGCTTACTGATCAGCTGGGATGCTCTGCTGTGCAC 4689  
Db 4724 GGAAGTGTGCTGCGACCCCAACAGCTTACTGATCAGCTGGGATGCTCTGCTGTGCAC 4783  
Qy 4690 AGTGAGATATTACAGATCACTTACGAGAAACAGAGAGAAATAGCCCTGTCAGGAGTT 4749  
Db 4784 AGTGAGATATTACAGATCACTTACGAGAAACAGAGAGAAATAGCCCTGTCAGGAGTT 4843  
Qy 4750 CACTGTGCTGGAGCAAGTCTACAGCTTACCATCAGCGGCTTAAACCTGGAGTTGATTA 4809  
Db 4844 CACTGTGCTGGAGCAAGTCTACAGCTTACCATCAGCGGCTTAAACCTGGAGTTGATTA 4903  
Qy 4810 TACCATCACTGTGTATGCTTCACTGGCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT 4869  
Db 4904 TACCATCACTGTGTATGCTTCACTGGCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT 4963  
Qy 4870 TTCCATTAATTACCGAACAGAAATTGACAAACCATCCAGATGCAAGTGACCGATGTTCA 4929  
Db 4964 TTCCATTAATTACCGAACAGAAATTGACAAACCATCCAGATGCAAGTGACCGATGTTCA 5023  
Qy 4930 GGACAACAGCATTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGGTTACAGAGT 4989  
Db 5024 GGACAACAGCATTAGTGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGGTTACAGAGT 5083  
Qy 4990 AACCCACTCTCCAAAATGGACAGGACCAACAAAATTAACCTGCAAGTCCAGATCA 5049  
Db 5084 AACCCACTCTCCAAAATGGACAGGACCAACAAAATTAACCTGCAAGTCCAGATCA 5143  
Qy 5050 AACAGAAATGACTATTGAAGGCTTGCAGCCCACTGGAGATGTTGTTAGTGTCTATGC 5109  
Db 5144 AACAGAAATGACTATTGAAGGCTTGCAGCCCACTGGAGATGTTGTTAGTGTCTATGC 5203  
Qy 5110 TCAGATCCAAAGGAGAGAGTTCAGGCTCTGTTTCAAGTTCCTGTTACTGGTTACAGAGT 5169  
Db 5204 TCAGATCCAAAGGAGAGAGTTCAGGCTCTGTTTCAAGTTCCTGTTACTGGTTACAGAGT 5263  
Qy 5170 CCTAAAGGACTGGCAATTCAGTGTGATGTCGATTCGATTCGATTCGATTCGATTCGATTC 5229  
Db 5264 CCTAAAGGACTGGCAATTCAGTGTGATGTCGATTCGATTCGATTCGATTCGATTCGATTC 5323

Qy 5230 CCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGATGGAATCCA 5289  
Db 5324 CCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGATGGAATCCA 5383  
Qy 5290 TGAGCTATTCCCTGCGACCTGATGGTGAAGAACAACACTGAGAGCTGCAAGGCTCAGACC 5349  
Db 5384 TGAGCTATTCCCTGCGACCTGATGGTGAAGAACAACACTGAGAGCTGCAAGGCTCAGACC 5443  
Qy 5350 GGTTCTGAGTACACAGTCACTGCTGCTTCCCTGCAAGATATGGAAGACGAGCCCT 5409  
Db 5444 GGTTCTGAGTACACAGTCACTGCTGCTTCCCTGCAAGATATGGAAGACGAGCCCT 5503  
Qy 5410 GATTGAAACCCAGTCCACAGCTATTCTGCAACCACTGACCTGAAAGTTCACTCAGGTCA 5469  
Db 5504 GATTGAAACCCAGTCCACAGCTATTCTGCAACCACTGACCTGAAAGTTCACTCAGGTCA 5563  
Qy 5470 ACCACAAGCCTGAGCGCCAGTGGACACCAACCAATGTTGAGCTCACTGGATATCGAGT 5529  
Db 5564 ACCACAAGCCTGAGCGCCAGTGGACACCAACCAATGTTGAGCTCACTGGATATCGAGT 5623  
Qy 5530 GCGGTGACCCCAAGGAGAACCGGACCAATGAAGAAATCAACCTTCTCTGACAG 5589  
Db 5624 GCGGTGACCCCAAGGAGAACCGGACCAATGAAGAAATCAACCTTCTCTGACAG 5683  
Qy 5590 CTCACTCCGTGTTGATCAGGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGC 5649  
Db 5684 CTCACTCCGTGTTGATCAGGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGC 5743  
Qy 5650 TCTTAAGGACATTTGACAGCAGACCACTCAGGAGTGTGACCACTCTGGAGAAATGT 5709  
Db 5744 TCTTAAGGACATTTGACAGCAGACCACTCAGGAGTGTGACCACTCTGGAGAAATGT 5803  
Qy 5710 CAGCCCAACCAAGAGGGCTCTGTGACAGATGCTACTTGAGACCAACCATCACTTAGCTG 5769  
Db 5804 CAGCCCAACCAAGAGGGCTCTGTGACAGATGCTACTTGAGACCAACCATCACTTAGCTG 5863  
Qy 5770 GAGAACCAAGATGAGACGATCACTGGGCTTCAAGTTGATGCGGTTCCAGCAATGGCCA 5829  
Db 5864 GAGAACCAAGATGAGACGATCACTGGGCTTCAAGTTGATGCGGTTCCAGCAATGGCCA 5923  
Qy 5830 GACTCAATCCAGAGAACCATCAAGCCAGATGTCAGAGCTACACCATCACTGGGTTTACA 5889  
Db 5924 GACTCAATCCAGAGAACCATCAAGCCAGATGTCAGAGCTACACCATCACTGGGTTTACA 5983  
Qy 5890 ACCAGGACTGACTACAGATCTACTGACACTTGAATGACCAATGCTCGAGCTCCCC 5949  
Db 5984 ACCAGGACTGACTACAGATCTACTGACACTTGAATGACCAATGCTCGAGCTCCCC 6043  
Qy 5950 TGTGGTCACTGAGCGCTTCACTGCAATTCATGACCATCCAACTGCGTTTCTGGCCAC 6009  
Db 6044 TGTGGTCACTGAGCGCTTCACTGCAATTCATGACCATCCAACTGCGTTTCTGGCCAC 6103  
Qy 6010 CACACCAATTCCTGCTGATATGAGGAGCGGCACTGCGGAGATTAACGGCTACAT 6069  
Db 6104 CACACCAATTCCTGCTGATATGAGGAGCGGCACTGCGGAGATTAACGGCTACAT 6163  
Qy 6070 CATCAAGTATGAGAGCTGCGGCTCTCCCAAGAGAGTGTCTCCGCGCCCGCCCTGG 6129  
Db 6164 CATCAAGTATGAGAGCTGCGGCTCTCCCAAGAGAGTGTCTCCGCGCCCGCCCTGG 6223  
Qy 6130 TGTCAAGAGGCTACTATTACTGCGCTGGAACCGGGAACCGAATATACAAATTAATGTCA 6189  
Db 6224 TGTCAAGAGGCTACTATTACTGCGCTGGAACCGGGAACCGAATATACAAATTAATGTCA 6283  
Qy 6190 TGCCTCTGAAAGATTAATCAGAGAGCGGCGCTGATTGGAAGGAAAACAGACAGAGCT 6249  
Db 6284 TGCCTCTGAAAGATTAATCAGAGAGCGGCGCTGATTGGAAGGAAAACAGACAGAGCT 6343  
Qy 6250 TCCCAACTGGTAACCTTCCACACCCCAATCTTATGAGACAGAGATCTTGGATGTTCC 6309  
Db 6344 TCCCAACTGGTAACCTTCCACACCCCAATCTTATGAGACAGAGATCTTGGATGTTCC 6403

2Y	6310	TTCACAGTTCAAAGAGACCCCTTTTGTCACCCACCCCTGGGTATGACACTGGAAATGGTAT	6369
2b	6404	TTCACAGTTCAAAGAGACCCCTTTGCGTACCCACCCCTGGGTATGACACTGGAAATGGTAT	6463
2Y	6370	TCAGCTTCCTGGCACATTCCTGGTCAGCAACCCAGTGTGTGGGCAACAAATGATCTTTGAGGA	6429
2b	6464	TCAGCTTCCTGGCACATTCCTGGTCAGCAACCCAGTGTGTGGGCAACAAATGATCTTTGAGGA	6523
2Y	6430	ACATGGTTTATGGCGGACACACCGCGCCACAACCGGCCACCCCCATTAAGGCATAGGCCAAG	6489
2b	6524	ACATGGTTTATGGCGGACACACCGCGCCACAACCGGCCACCCCCATTAAGGCATAGGCCAAG	6583
2Y	6490	ACCATACCCGCGCAATGTAGGACAAGAAGCTCTCTCTCAGACAACCATCTCATAGGCCCC	6549
2b	6584	ACCATACCCGCGCAATGTAGGACAAGAAGCTCTCTCTCAGACAACCATCTCATAGGCCCC	6643
2Y	6550	ATTCAGGACACTTCGTAGTACATATTTCAATGTCACTCTGTGTGGCACTGATGAAGAACC	6609
2b	6644	ATTCAGGACACTTCGTAGTACATATTTCAATGTCACTCTGTGTGGCACTGATGAAGAACC	6703
2Y	6610	CTTACAGTTACAGGGTTCCTGGAACTTCTACCACTGGCCACTCTCAGACGGCTCACAGAGG	6669
2b	6704	CTTACAGTTACAGGGTTCCTGGAACTTCTACCACTGGCCACTCTCAGACGGCTCACAGAGG	6763
2Y	6670	TGCCACCTCAACATCATAGTGGAGGCACATGAAGAAGCACGACAGGCGATAAGGTTCCGGA	6729
2b	6764	TGCCACCTCAACATCATAGTGGAGGCACATGAAGAAGCACGACAGGCGATAAGGTTCCGGA	6823
2Y	6730	AGAGGTGTGTACCGTGGGCAACTCTGTCAACGAAGGCTTGAAACAACCTTACGATGACTC	6789
2b	6824	AGAGGTGTGTACCGTGGGCAACTCTGTCAACGAAGGCTTGAAACAACCTTACGATGACTC	6883
2Y	6790	GTGCTTTGACCCCTCACACAGTGTCCCATATATGCGTTGGAGATGAGTGGGAAACGAATGTC	6849
2b	6884	GTGCTTTGACCCCTCACACAGTGTCCCATATATGCGTTGGAGATGAGTGGGAAACGAATGTC	6943
2Y	6850	TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTTAGGCTTTGGAAAGTGTCATTTCCAGATG	6909
2b	6944	TGAATCAGGCTTTAAACTGTGTGCCAGTGCTTTAGGCTTTGGAAAGTGTCATTTCCAGATG	7003
2Y	6910	TGATTCATCTAGATGTGCCATCACAAATGGTGTGAACTACAAGATTGGAGAGTGGGA	6969
2b	7004	TGATTCATCTAGATGTGCCATCACAAATGGTGTGAACTACAAGATTGGAGAGTGGGA	7063
2Y	6970	CCGTCAGGGAGAAAATGGCCAGATGATGAGCTGCACATGTCTTTGGAAACGGAAAAGGAGA	7029
2b	7064	CCGTCAGGGAGAAAATGGCCAGATGATGAGCTGCACATGTCTTTGGAAACGGAAAAGGAGA	7123
2Y	7030	ATTCAAGTGTGACCTTCATGAGGCAACGTGTTACGATGATGGGAAGACATACACAGTAGG	7089
2b	7124	ATTCAAGTGTGACCTTCATGAGGCAACGTGTTACGATGATGGGAAGACATACACAGTAGG	7183
2Y	7090	AGAACAGTGGCAGAAAGAAATATCTCGGTGCCATTTGCTCTCTGCACATGCTTTTCGAGGCCA	7149
2b	7184	AGAACAGTGGCAGAAAGAAATATCTCGGTGCCATTTGCTCTCTGCACATGCTTTTCGAGGCCA	7243
2Y	7150	GCGGGCTGGCGTGTGACAACTGCGCGCAGACCTGGGGGTGAACCCAGTCCCCGAAGGCAC	7209
2b	7244	GCGGGCTGGCGTGTGACAACTGCGCGCAGACCTGGGGGTGAACCCAGTCCCCGAAGGCAC	7303
2Y	7210	TACTGGCCAGTCTCAACACAGATTTCTCAGAGATACCATCAGAGAACAACTAAATGT	7269
2b	7304	TACTGGCCAGTCTCAACACAGATTTCTCAGAGATACCATCAGAGAACAACTAAATGT	7363
2Y	7270	TAAATTGCCAATTGAGTGTCTTACGCTTTTAGATGTACAGGCTGACAGAGAAATTTCCCG	7329
2b	7364	TAAATTGCCAATTGAGTGTCTTACGCTTTTAGATGTACAGGCTGACAGAGAAATTTCCCG	7423
2Y	7330	AGAGTAAATCATCTTTTCCAACTCAGAGGACACAGCATGTCTCTCTGCCAGATCCATCTA	7389
2b	7424	AGAGTAAATCATCTTTTCCAACTCAGAGGACACAGCATGTCTCTCTGCCAGATCCATCTA	7483
2Y	7390	AACTGGAGTGAATGTAGCAGACCCAGCTTAGAGTTCCTTTCTTTTCTTAAGCCCTTTGC	7449

Db	7484	A	ACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTTCTTCTTTCTTTCTTAAGCCCTTTGC	7543
Qy	7450	T	CTGGAGGAGAGTTCTCCAGCTTCAGCTTCAACTCAGCTTCTCCAAAGCATCACCTCGGGA	7509
Db	7544	T	CTGGAGGAGAGTTCTCCAGCTTCAGCTTCAACTCAGCTTCTCCAAAGCATCACCTCGGGA	7603
Qy	7510	G	TTTCTCTGAGGGTTTCTCTAATAATGAGGGCTGCACATTCGCTGTGTTCTGTTTGAAGTAT	7569
Db	7604	G	TTTCTCTGAGGGTTTCTCTAATAATGAGGGCTGCACATTCGCTGTGTTCTGTTTGAAGTAT	7663
Qy	7570	T	CAATACCGCTCAGTATTTTAAATGAAGTGATTCTAAGATTTCGTTTGGGATCAATAGGA	7629
Db	7664	T	CAATACCGCTCAGTATTTTAAATGAAGTGATTCTAAGATTTCGTTTGGGATCAATAGGA	7723
Qy	7630	A	AGCATATGCAGGCCAACACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAAAGT	7689
Db	7724	A	AGCATATGCAGGCCAACACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAAAGT	7783
Qy	7690	A	GAAGAGTGCACCAACACACTTCTGCTTTCACTTAAGTGTGCGCCGCATCTGTAGGA	7749
Db	7784	A	GAAGAGTGCACCAACACACTTCTGCTTTCACTTAAGTGTGCGCCGCATCTGTAGGA	7843
Qy	7750	A	CAAGCATGATCTTGTTACTGTGATATTTTAAATATCCACAGTACT	7795
Db	7844	A	CAAGCATGATCTTGTTACTGTGATATTTTAAATATCCACAGTACT	7889
RESULT 3				
AAI57804				
XX	ID	AAI57804	standard; cdna; 8230 BP.	
XX	AC	AAI57804;		
XX	DT	22-OCT-2001	(first entry)	
XX	DE	Human polynucleotide	SEQ ID NO 7.	
XX	KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
XX	KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
XX	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
XX	KW	anoychotic lateral sclerosis; Shy-Drager syndrome; chemotactic;		
XX	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
XX	KW	leukaemia; ss.		
OS		Homo sapiens.		
XX	FN	WO20015312-A1.		
XX	PD	26-JUL-2001.		
XX	PF	26-DEC-2000; 2000WO-US034263.		
XX	PR	23-DEC-1999; 99US-00471275.		
XX	PR	21-JAN-2000; 2000US-00488725.		
XX	PR	25-APR-2000; 2000US-00552317.		
XX	PR	20-JUN-2000; 2000US-00598042.		
XX	PR	19-JUL-2000; 2000US-00620312.		
XX	PR	03-AUG-2000; 2000US-00653450.		
XX	PR	14-SEP-2000; 2000US-00662191.		
XX	PR	19-OCT-2000; 2000US-00693036.		
XX	PR	29-NOV-2000; 2000US-00727344.		
XX	FA	(HYSE-) HYSEO INC.		









6190 TGCCCTGAAGAAATATCATGAGAGAGCGAGCCCTGTGATTTGGAAGGAAAAAGACAGACGAGCT 6249  
 b 6284 TGCCCTGAAGAAATATCATGAGAGAGCGAGCCCTGTGATTTGGAAGGAAAAAGACAGACGAGCT 6343  
 y 6250 TCCCCAACTGGTAAACCTTCCACACCCCAATCTTCAATGGACAGAGATCTTGGATGTTCC 6309  
 b 6344 TCCCCAACTGGTAAACCTTCCACACCCCAATCTTCAATGGACAGAGATCTTGGATGTTCC 6403  
 y 6310 TTCCACAGTTCAAAAGACCCCTTTTGTGTCACCCACCCCTGGTATGACATGGAAATGGTAT 6369  
 b 6404 TTCCACAGTTCAAAAGACCCCTTTTGTGTCACCCACCCCTGGTATGACATGGAAATGGTAT 6463  
 y 6370 TCAGCTTCTCTGCACTTCTGCTGAGCAACCCCAATCTTCAATGGACAGAGATCTTGGATGTTCC 6429  
 b 6464 TCAGCTTCTCTGCACTTCTGCTGAGCAACCCCAATCTTCAATGGACAGAGATCTTGGATGTTCC 6523  
 y 6430 ACATGGTTTTAGCGGACACACCCGCCCAACCGGCCATTAAGGCATAGGCCAAG 6489  
 b 6524 ACATGGTTTTAGCGGACACACCCGCCCAACCGGCCATTAAGGCATAGGCCAAG 6583  
 y 6490 ACCATACCCCGCAATGT----- 6507  
 b 6584 ACCATACCCCGCAATGTAGGTGAGGAAATCCAAATGGTCAATCCCAAGGAGATGT 6643  
 y 6508 -----AGGACAGA 6516  
 b 6644 AGCATATCACCTGTACCCACACCGTCCGGGGCTCAATCCAAATGGCTCTACAGGACAGA 6703  
 y 6517 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCAGGACATCTCTGAGTACATCAT 6576  
 b 6704 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCAGGACATCTCTGAGTACATCAT 6763  
 y 6577 TTCACTGTCATCTGTTGGCACTGATGAGAACCTTACAGTTCAGGGTTCCTGGAACCTC 6636  
 b 6764 TTCACTGTCATCTGTTGGCACTGATGAGAACCTTACAGTTCAGGGTTCCTGGAACCTC 6823  
 y 6637 TACCAGTGCATCTGACAGGCTTACAGAGGTCACCTACACATCATAGTGGAGGC 6696  
 b 6824 TACCAGTGCATCTGACAGGCTTACAGAGGTCACCTACACATCATAGTGGAGGC 6883  
 y 6697 ACTGAAAGACACGAGGCAATAGGTTCCGGAAGAGGTTGTTACCGTGGGCACTCTGT 6756  
 b 6884 ACTGAAAGACACGAGGCAATAGGTTCCGGAAGAGGTTGTTACCGTGGGCACTCTGT 6943  
 y 6757 CAACGAGGCTTTGAACCAACCTACGAGTACCTGCTTTGACCCCTACACAGTGTCCCA 6816  
 b 6944 CAACGAGGCTTTGAACCAACCTACGAGTACCTGCTTTGACCCCTACACAGTGTCCCA 7003  
 y 6817 TTATGCCGTGGAGATGAGTGGAGGAAATGCTGAATCAGGCTTTAACTGTTGGCA 6876  
 b 7004 TTATGCCGTGGAGATGAGTGGAGGAAATGCTGAATCAGGCTTTAACTGTTGGCA 7063  
 y 6877 GTGCTTAGGCTTTGGAAGTGGTCAATTCAGATGCTGATTCATCTAGATGGTGCATGACAA 6936  
 b 7064 GTGCTTAGGCTTTGGAAGTGGTCAATTCAGATGCTGATTCATCTAGATGGTGCATGACAA 7123  
 y 6937 TGGTGTGAATCAAGATTTGGAAGTGGAGGCGTTCAGGAGAAAATGCCAGATGAT 6996  
 b 7124 TGGTGTGAATCAAGATTTGGAAGTGGAGGCGTTCAGGAGAAAATGCCAGATGAT 7183  
 y 6997 GAGCTGCATGCTTTGGGAAACGGAAGGAGAAATTCAGTGTGACCTCATGAGGCAAC 7056  
 b 7184 GAGCTGCATGCTTTGGGAAACGGAAGGAGAAATTCAGTGTGACCTCATGAGGCAAC 7243  
 y 7057 GTGTTACGATGATGGGAAGACATACCACTAGAGAACAGTGGCAGAGGAAATATCTCGG 7116  
 b 7244 GTGTTACGATGATGGGAAGACATACCACTAGAGAACAGTGGCAGAGGAAATATCTCGG 7303  
 y 7117 TGCCTATTGCTCTGTCACATGCTTTGGAGCCAGGCGCTGTCGCTGTGACAACTGCGG 7176  
 b 7304 TGCCTATTGCTCTGTCACATGCTTTGGAGCCAGGCGCTGTCGCTGTGACAACTGCGG 7363  
 y 7177 CAGACCTGGGGTGAACCCAGCTCCGAGGCACTAGTGGCCAGTCTCTACAAACCACTATTC 7236

7364 CAGACCTGGGGTGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTACACCACTATTC 7423  
 y 7237 TCAGAGATACCATCAGAGAACAACTAATGTTAATTTGCCAAATTTGAGTGTCTTATGCC 7296  
 b 7424 TCAGAGATACCATCAGAGAACAACTAATGTTAATTTGCCAAATTTGAGTGTCTTATGCC 7483  
 y 7297 TTTAGATGTACAGCTGACAGAGAAATTTCCCGAGAGTAAATCATCTTTCCATCCAGAG 7356  
 b 7484 TTTAGATGTACAGCTGACAGAGAAATTTCCCGAGAGTAAATCATCTTTCCATCCAGAG 7543  
 y 7357 GAAACAGCATGCTCTCTGCGCAAGATCCATCTAAACTGGAGTATCTTAGCAGACCCAGC 7416  
 b 7544 GAAACAGCATGCTCTCTGCGCAAGATCCATCTAAACTGGAGTATCTTAGCAGACCCAGC 7603  
 y 7417 TTAGAGTTCTTCTTCTTCTTAAAGCCCTTCTGCGAGGAGTCTTCTGAGGTTCTCTATAAATGA 7476  
 b 7604 TTAGAGTTCTTCTTCTTCTTAAAGCCCTTCTGCGAGGAGTCTTCTGAGGTTCTCTATAAATGA 7663  
 y 7477 CAACCTCACAGCTTCTCCAGCATCACCTGGGAGTCTTCTGAGGTTCTCTATAAATGA 7536  
 b 7664 CAACCTCACAGCTTCTCCAGCATCACCTGGGAGTCTTCTGAGGTTCTCTATAAATGA 7723  
 y 7537 GGCTGACATTCCTGTTCTGCTTGAAGTATTTCAATACCGCTCAGTATTTTAAATGAA 7596  
 b 7724 GGCTGACATTCCTGTTCTGCTTGAAGTATTTCAATACCGCTCAGTATTTTAAATGAA 7783  
 y 7597 GTGATTTCAAGATTTGTTGGGATCAATAGGAAAGCATATGAGCAACCAAGATGCAA 7656  
 b 7784 GTGATTTCAAGATTTGTTGGGATCAATAGGAAAGCATATGAGCAACCAAGATGCAA 7843  
 y 7657 ATGTTTGAATGATATGACCAAAATTTTAAAGTGAAGAAAGTACCCAAACACTTCTGCTT 7716  
 b 7844 ATGTTTGAATGATATGACCAAAATTTTAAAGTGAAGAAAGTACCCAAACACTTCTGCTT 7903  
 y 7717 TCATTTAAGTGTCTGCGCCGCAATCTGAGGAAAGCATATGAGCAACCAAGATGCTTGTACTGTGATAT 7776  
 b 7904 TCATTTAAGTGTCTGCGCCGCAATCTGAGGAAAGCATATGAGCAACCAAGATGCTTGTACTGTGATAT 7963  
 y 7777 TTTAAATATCCACAGTACT 7795  
 b 7964 TTTAAATATCCACAGTACT 7982

RESULT 4  
 AA157802  
 ID AA157802 standard; cDNA; 8062 BP.  
 XX AC AA157802;  
 XX AC AA157802;  
 XX DT 22-OCT-2001 (first entry)  
 XX XX Human polynucleotide SEQ ID NO 5.  
 DE Human; nontopic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX OS Homo sapiens.  
 XX OS Homo sapiens.  
 XX PN WO200153312-A1.  
 XX XX 26-JUL-2001.  
 XX PF 26-DEC-2000; 2000WO-US034263.  
 XX PR 23-DEC-1999; 99US-00471275.  
 XX PR 21-JAN-2000; 2000US-00489725.  
 XX PR 25-APR-2000; 2000US-00552317.  
 XX PR 20-JUN-2000; 2000US-00598042.



1570 CACTTCTGAGGCGAGAGACAACATGAAGTGGTGGGACCAACAGAACTATGATGC 1629  
 1664 CACTTCTGAGGCGAGAGACAACATGAAGTGGTGGGACCAACAGAACTATGATGC 1723  
 1630 CGACGAGAGATTTGGGTTCTGCCCCATGCTGCCACAGAGGAATCTGCAACCAATGA 1689  
 1724 CGACGAGAGATTTGGGTTCTGCCCCATGCTGCCACAGAGGAATCTGCAACCAATGA 1783  
 1690 AGGGGTCACTGATCCGTCATTTGGAGATCAGTGGGATGAAGCAGATGATGGGTCAATGAT 1749  
 1784 AGGGGTCACTGATCCGTCATTTGGAGATCAGTGGGATGAAGCAGATGATGGGTCAATGAT 1843  
 1750 GAGGTCACTGATTTGGGAAATGGTCTGGGAAATGGGATGATGATGATGATGATGATGAT 1809  
 1844 GAGGTCACTGATTTGGGAAATGGTCTGGGAAATGGGATGATGATGATGATGATGATGAT 1903  
 1810 TCAGATCAGTGCATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1869  
 1904 TCAGATCAGTGCATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1963  
 1870 TCATGAGAGAGGCAATGCTGAACTGTATGATGATGATGATGATGATGATGATGATGATGAT 1929  
 1964 TCATGAGAGAGGCAATGCTGAACTGTATGATGATGATGATGATGATGATGATGATGATGAT 2023  
 1930 GTGTGATCCGTCGACCAATGCCAGATTCAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1989  
 2024 GTGTGATCCGTCGACCAATGCCAGATTCAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 2083  
 1990 TCCATGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2049  
 2084 TCCATGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2143  
 2050 TGGGAGTGGCATGTCACCTTTACAGACCTTCCAGCTCAAGTGGTCTGTCGAGT 2109  
 2144 TGGGAGTGGCATGTCACCTTTACAGACCTTCCAGCTCAAGTGGTCTGTCGAGT 2203  
 2110 ATTATTCATGAGATCCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2169  
 2204 ATTATTCATGAGATCCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2263  
 2170 GCATCTCAGATTTCCAGTACATTTCCAGTACATTTCCAGTACATTTCCAGTACATTTCCAGTACAT 2229  
 2264 GCATCTCAGATTTCCAGTACATTTCCAGTACATTTCCAGTACATTTCCAGTACATTTCCAGTACAT 2323  
 2230 GAAGGAAGTACCATACAGGCGCATTTAACTCTACACCATCAAGGCTCAAGCCTGG 2289  
 2324 GAAGGAAGTACCATACAGGCGCATTTAACTCTACACCATCAAGGCTCAAGCCTGG 2383  
 2290 TGTGGTATACAGGCGCATTTATCAGATCCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 2349  
 2384 TGTGGTATACAGGCGCATTTATCAGATCCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 2443  
 2350 CTATTGACTTACACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2409  
 2444 CTATTGACTTACACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2503  
 2410 GAGGACTCCCTTTCT 2469  
 2504 GAGGACTCCCTTTCT 2563  
 2470 TAGCTTTGGTCT 2529  
 2564 TAGCTTTGGTCT 2623  
 2530 TAGCTTTGGTCT 2589  
 2624 TAGCTTTGGTCT 2683  
 2590 TGTGACATCCCTGACCT 2649  
 2684 TGTGACATCCCTGACCT 2743  
 2650 TGAGGATGGGAGCAGAGTTTGTATCT 2709

2744 TGAGGATGGGAGCAGAGTTTGTATCT 2803  
 2710 TCCTGACCCGACTGTGACCAAGTGTGATGACACTCAATGTTGTTGCTGCTGAGCAGAC 2769  
 2804 TCCTGACCCGACTGTGACCAAGTGTGATGACACTCAATGTTGTTGCTGCTGAGCAGAC 2863  
 2770 CCAGGCTCCCATCACAGGCTACAGAAATAGTCTATTCGCCATCAGTAGAAGGTAGCAGCAC 2829  
 2864 CCAGGCTCCCATCACAGGCTACAGAAATAGTCTATTCGCCATCAGTAGAAGGTAGCAGCAC 2923  
 2830 AGAATCAACTCTCTGAAACTGCAACTCCGTCAGCTCAGTGAATGCAACTGCTGCTGCT 2889  
 2924 AGAATCAACTCTCTGAAACTGCAACTCCGTCAGCTCAGTGAATGCAACTGCTGCTGCT 2983  
 2890 TCAGTATAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACACCTGCTGCTGCT 2949  
 2984 TCAGTATAACATCACTATCTATGCTGTGGAAGAAATCAAGAAAGTACACCTGCTGCTGCT 3043  
 2950 TCACAAAGAAACCACTGGCAACCCCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 3009  
 3044 TCACAAAGAAACCACTGGCAACCCCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 3103  
 3010 GTTGTGGAAGTACAGACGCTGAAGTACACATCATGTGACACCCGCTGAGAGTGCAGT 3069  
 3104 GTTGTGGAAGTACAGACGCTGAAGTACACATCATGTGACACCCGCTGAGAGTGCAGT 3163  
 3070 GACCGGTACTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3129  
 3164 GACCGGTACTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3223  
 3130 GCCATCAGCAGAGAACCTTTTCAGAGTACCGGCTGCTCCCTGGGCTCACTTATTA 3189  
 3224 GCCATCAGCAGAGAACCTTTTCAGAGTACCGGCTGCTCCCTGGGCTCACTTATTA 3283  
 3190 CTTCAAAGTCTTTGCACTGAGCCTTGGGAGGAGCAAGCTCTGACTGCTCAACAGAC 3249  
 3284 CTTCAAAGTCTTTGCACTGAGCCTTGGGAGGAGCAAGCTCTGACTGCTCAACAGAC 3343  
 3250 AACCAACTGGATGCTCCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3309  
 3344 AACCAACTGGATGCTCCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3403  
 3310 GGTGAGATGAGTCTCCACTCGGCGCCAGATAAAGGATACCGACTGACCGTGGGCTTTAC 3369  
 3404 GGTGAGATGAGTCTCCACTCGGCGCCAGATAAAGGATACCGACTGACCGTGGGCTTTAC 3463  
 3370 CCGAAGAGCAGCCGAGCAGTACATGTTGGTCCCTCTGCTCCAGTACCGACTGAG 3429  
 3464 CCGAAGAGCAGCCGAGCAGTACATGTTGGTCCCTCTGCTCCAGTACCGACTGAG 3523  
 3430 GAATCTGAGCTGCTGAGTACACCGTATCCCTCGTGGCCATTAAGGGCAACCAAGA 3489  
 3524 GAATCTGAGCTGCTGAGTACACCGTATCCCTCGTGGCCATTAAGGGCAACCAAGA 3583  
 3490 GAGCCCCAAGCCACTGGAGTCTTTACCACTGCGAGCTGGGAGTCTATTCACCTTA 3549  
 3584 GAGCCCCAAGCCACTGGAGTCTTTACCACTGCGAGCTGGGAGTCTATTCACCTTA 3643  
 3550 CAACACCGAGTGTGAGACCACTTGTGATCAGATGAGCCTGCTGCTCCAGAAATGG 3609  
 3644 CAACACCGAGTGTGAGACCACTTGTGATCAGATGAGCCTGCTGCTCCAGAAATGG 3703  
 3610 TTTTAACTGGGTGTACGACCAAGCCAGGAGGAGGACCAACGAGAGTGTGCTTACCA 3669  
 3704 TTTTAACTGGGTGTACGACCAAGCCAGGAGGAGGAGGACCAACGAGAGTGTGCTTACCA 3763  
 3670 CTGAGGAGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3729  
 3764 CTGAGGAGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3823  
 3730 AGTCTCTGAGAGATGAGCAGGAAAGAGATGCGCCAAATTTGTAAACAAAGGTGTGACACCAAT 3789





5950	2Y	TTGTGTCATCGACGCCTCCACTGCCATTGATGACCAATCCAACTCTGGTTTCTGTGCCAC	6009
6044	Db	TTGTGTCATCGACGCCTCCACTGCCATTGATGACCAATCCAACTCTGGTTTCTGTGCCAC	6103
6010	2Y	CACACCCAAATTCCTTGCTGGTATCATGGACGCGCCACGTCGCCAGATTACCGCTTACAT	6069
6104	Db	CACACCCAAATTCCTTGCTGGTATCATGGACGCGCCACGTCGCCAGATTACCGCTTACAT	6163
6070	2Y	CATCAAGTATGAGAACCTGGGTCTCTCTCCACAGAAAGTGGTCCCTCGGCCCCCGCCCTGG	6129
6164	Db	CATCAAGTATGAGAACCTGGGTCTCTCTCCACAGAAAGTGGTCCCTCGGCCCCCGCCCTGG	6223
6130	2Y	TGTCACAGAGCTACTATTACTTCGCTGGAAACCGGGAAACCGAATATACAAATTTATGTCAT	6189
6224	Db	TGTCACAGAGCTACTATTCTCTGGCTGGAAACCGGGAAACCGAATATACAAATTTATGTCAT	6283
6190	2Y	TGCCCTGAAGAAATAATCAGAAGACGAGCCCTGATTGGAAGAAAAAGACACAGAGCT	6249
6284	Db	TGCCCTGAAGAAATAATCAGAAGACGAGCCCTGATTGGAAGAAAAAG-----	6332
6250	2Y	TCCCAACTGGTAAACCTTCTCCACACCCCAATCTTCATGGACACAGATCTTGATGTTCC	6309
6333	Db	-----	6332
6310	2Y	TTCCACAGTTCAAAAGACCCCTTTGTCACCCACCCCTGGGTATGACACTGGAAATGGTAT	6369
6333	Db	-----ACAGTTCAAAAGACCCCTTTGTCACCCACCCCTGGGTATGACACTGGAAATGGTAT	6388
6370	2Y	TCAGCTTCCTGGCACTTCCTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA	6429
6389	Db	TCAGCTTCCTGGCACTTCCTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA	6448
6430	2Y	ACATGTTTTAGCGGACACACACGCGCCACAAACGGCCACCCCAATAGGCATAGGCCAAG	6489
6449	Db	ACATGTTTTAGCGGACACACACGCGCCACAAACGGCCACCCCAATAGGCATAGGCCAAG	6508
6490	2Y	ACCATACCGCCGAAATGTAGGAACAAGATCTCTCTCAGACAAACCATCTCATTGGGCCCC	6549
6509	Db	ACCATACCGCCGAAATGTAGGAACAAGATCTCTCTCAGACAAACCATCTCATTGGGCCCC	6568
6550	2Y	ATTCCAGGACATCTTCAGTACATATTTTCATCTCATCTCGTTGGCACTGATCAAGAAC	6609
6569	Db	ATTCCAGGACATCTTCAGTACATATTTTCATCTCATCTCGTTGGCACTGATCAAGAAC	6628
6610	2Y	CTTTACAGTTTCAGGGTTCTGGAACTTCTACAGTGGCACTCTCAGACGGCTCACACAGG	6669
6629	Db	CTTTACAGTTTCAGGGTTCTGGAACTTCTACAGTGGCACTCTCAGACGGCTCACACAGG	6688
6670	2Y	TGCCACCTACAAATCATAGTGGAGGCACTGAAGAACCAGCAGAGGCATTAAGTTGGGA	6729
6689	Db	TGCCACCTACAAATCATAGTGGAGGCACTGAAGAACCAGCAGAGGCATTAAGTTGGGA	6748
6730	2Y	AGAGGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTTGAAACCAACCTACGGATGACTC	6789
6749	Db	AGAGGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTTGAAACCAACCTACGGATGACTC	6808
6790	QY	GTGCTTTGACCCCTACACAGTGTCCATTTATGCGGTGGAGATGAGTGGGAAACGAATGTC	6849
6809	Db	GTGCTTTGACCCCTACACAGTGTTCCTATTTAGCCCTTTGGAGTGGAGATGAGTGGGAAACGAATGTC	6868
6850	QY	TGAATCAGGCTTTAAACTGTGTGGCAGTGCTTTAGGCTTTTGGAAAGTGGTTCATTTTCAGATG	6909
6869	Db	TGAATCAGGCTTTAAACTGTGTGGCAGTGCTTTAGGCTTTTGGAAAGTGGTTCATTTTCAGATG	6928
6910	QY	TGATTCATCTAGATGGTGCCATGCAATGGTGTGAATACAAATTGGAGAGAGTGGGA	6969
6929	Db	TGATTCATCTAGATGGTGCCATGCAATGGTGTGAATACAAATTGGAGAGAGTGGGA	6988
6970	QY	CCGTCAGGAGAAATGGCCAGATGATGAGCTGCATGTCCTTTGGGAAACGGAAAAAGAGA	7029
6989	Db	CCGTCAGGAGAAATGGCCAGATGATGAGCTGCATGTCCTTTGGGAAACGGAAAAAGAGA	7048
7030	QY	ATTCAAGTGTGACCCCTCATGAGGCAACGTTTACGATGATGGGAAGACATATCACCGTAGG	7089

7049	Db	ATTTCAAGTGTGACCCCTCATGAGGCAACGTTTACGATGATGGGAAGACATACACACGTAGG	7108
7090	QY	AGACAAGTGGCAGAAAGGAATATCTCGGTGCCATTTCTCCTCTGCACATGCTTTTCGAGGCCA	7149
7109	Db	AGACAGTGGCAGAGAGGAATATCTCGGTGCCATTTGCTCTCTGCACATGCTTTTCGAGGCCA	7168
7150	QY	GGGGGGCTGGCGCTGTGACAACTGCCGAGACCTGGGGGTGAACCCAGTCCCCGAAAGGCAC	7209
7169	Db	GGGGGGCTGGCGCTGTGACAACTGCCGAGACCTGGGGGTGAACCCAGTCCCCGAAAGGCAC	7228
7210	QY	TACTGGCCAGTCTCTACAACAGTATTTCTCAGAGATACCATCAGAGAACAAACACTAATGT	7269
7229	Db	TACTGGCCAGTCTCTACAACAGTATTTCTCAGAGATACCATCAGAGAACAAACACTAATGT	7288
7270	QY	TAAATGCCCAATTGAGTGGTTTCAATGCTTTTAGATGTACAGGCTGAACAGAGAGATTTCCCG	7329
7289	Db	TAAATTGCCCAATTTGAGTGGTTTCAATGCTTTTAGATGTACAGGCTGAACAGAGAGATTTCCCG	7348
7330	QY	AGAGTAATAATCATCTTTCCAAATCCAGAGGAACAGCATGTCTCTGCCAAGATCCCATCTA	7389
7349	Db	AGAGTAATAATCATCTTTCCAAATCCAGAGGAACAGCATGTCTCTGCCAAGATCCCATCTA	7408
7390	QY	AACTGGAGTGATGTTAGCAGACCCAGCTTTAGAGTTCTTCTTTCTTAAAGCCCTTTGC	7449
7409	Db	AACTGGAGTGATGTTAGCAGACCCAGCTTTAGAGTTCTTCTTTCTTAAAGCCCTTTGC	7468
7450	QY	TCTGGAGGAAGTCTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATCACCCTGGGA	7509
7469	Db	TCTGGAGGAAGTCTCCAGCTTCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATCACCCTGGGA	7528
7510	QY	GTTTCTCTAGGGTTTTCTCATAAATGAGGGCTGCACATTCCTGTTCTGTTTCGAAAGTAT	7569
7529	Db	GTTTCTCTAGGGTTTTCTCATAAATGAGGGCTGCACATTCCTGTTCTGTTTCGAAAGTAT	7588
7570	QY	TCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGATTTGGTTGGGATCAATAGGA	7629
7589	Db	TCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGATTTGGTTGGGATCAATAGGA	7648
7630	QY	AAGCATATGCAGCAACCAAGATGCAAATGTTTGAATATGATATGACCAAAATTTTAAAGT	7689
7649	Db	AAGCATATGCAGCAACCAAGATGCAAATGTTTGAATATGATATGACCAAAATTTTAAAGT	7708
7690	QY	AGGAAGTCACCAAAACACTTCTGCTTTCACTTAAGTGCTCGGCCCGCAATACTGTAGGA	7749
7709	Db	AGGAAGTCACCAAAACACTTCTGCTTTCACTTAAGTGCTCGGCCCGCAATACTGTAGGA	7768
7750	QY	ACAACCATGATCTGTTTACTGATGATATTTTAAATATCCACAGTACT	7795
7769	Db	ACAACCATGATCTGTTTACTGATGATATTTTAAATATCCACAGTACT	7814
RESULT 5			
ADB31322			
ID	ADB31322 standard; cDNA; 7679 BP.		
AC	ADB31322;		
XX	ADB31322;		
DT	04-DEC-2003 (first entry)		
DE	Testosterone regulated prostate cancer gene #38.		
DE	Testosterone regulated prostate cancer gene #38.		
XX	Human; ss; prostate cancer; cancer; androgen; castration;		
XX	anti-androgenic drug; bicalutamide; casodex; testosterone.		
XX	Homo sapiens.		
XX	OS		
XX	US6506607-B1.		
XX	PN		
XX	14-JAN-2003.		
PD	23-DEC-1998; 98US-00220132.		
XX	PF		
XX	XX		



1621 GAGCTTTATCAAAATGGAGATTCAATGGAGAGATATGTCATGGTGCAGATACCAATG 1680  
2029 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACTTTACAGACCTATCCAAAG 2088  
1681 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACTTTACAGACCTATCCAAAG 1740  
2089 CTCAGTGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCACTCCACGCC 2148  
1741 CTCAGTGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCACTCCACGCC 1800  
2149 CATCCAGTGGAAATGCAACCAAGCCATCTCACATTTCCAAAGTACATTTCTCAGTGGAGACC 2208  
1801 CATCCAGTGGAAATGCAACCAAGCCATCTCACATTTCCAAAGTACATTTCTCAGTGGAGACC 1860  
2209 TAAAAATTTCTGTAGGCCGTGGAGAGACTACATACAGAGCCCACTTAAACTCTTACAC 2268  
1861 TAAAAATTTCTGTAGGCCGTGGAGAGACTACATACAGAGCCCACTTAAACTCTTACAC 1920  
2269 CATCAAGGCCCTGAAGCCTGGTGTGTATACGAGGCCAGCTCATCAGCATCCAGCAGTA 2328  
1921 CATCAAGGCCCTGAAGCCTGGTGTGTATACGAGGCCAGCTCATCAGCATCCAGCAGTA 1980  
2329 CGGCCACCAAGAGTCACTCGCTTTGACTTTCACCTACCAACCAAGCAGCAGCAGCAGCTGTGAC 2389  
1981 CGGCCACCAAGAGTCACTCGCTTTGACTTTCACCTACCAACCAAGCAGCAGCAGCAGCTGTGAC 2040  
2389 CAGCAACCGTGACAGGAG 2448  
2041 CAGCAACCGTGACAGGAG 2100  
2449 TGTGACCGAAATCACAGCCAGTAGCTTTGTGTCTCTCGGCTCTCAGCTTCCGACACCGT 2508  
2101 TGTGACCGAAATCACAGCCAGTAGCTTTGTGTCTCTCGGCTCTCAGCTTCCGACACCGT 2160  
2509 GTCCGAGTTCGGGTGGATATGAGCTGAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2568  
2161 GTCCGAGTTCGGGTGGATATGAGCTGAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2569 TCTTCCAAGCAGACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGCGCGGAAATACAT 2628  
2221 TCTTCCAAGCAGACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGCGCGGAAATACAT 2280  
2629 TGTAAATGTCTATCAGATATCTGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2688  
2281 TGTAAATGTCTATCAGATATCTGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2689 AACACAGCGCTGTATGCGCCCTCTGACCGAGTGTGGACCAAGTGTGATGACACCTCAAT 2748  
2341 AACACAGCGCTGTATGCGCCCTCTGACCGAGTGTGGACCAAGTGTGATGACACCTCAAT 2400  
2749 TGTGTTTGTGTCGAGCAGACCCAGGCTCCATACAGGGTACAGAAATAGTCTATTGCGC 2808  
2401 TGTGTTTGTGTCGAGCAGACCCAGGCTCCATACAGGGTACAGAAATAGTCTATTGCGC 2460  
2809 ATCAGTAGAAGGTAGCAGCAGAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCT 2868  
2461 ATCAGTAGAAGGTAGCAGCAGAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCT 2520  
2869 CAGTGAATTCGAACCTGGTGTTCAGTATAACATCACTATCTATGCTGTGGAAGAAATCA 2928  
2521 CAGTGAATTCGAACCTGGTGTTCAGTATAACATCACTATCTATGCTGTGGAAGAAATCA 2580  
2929 AGAAGTACACCTGTTGTCTATCAAGAGAAACCACTGGCAGCCCACTGATACAGT 2988  
2581 AGAAGTACACCTGTTGTCTATCAAGAGAAACCACTGGCAGCCCACTGATACAGT 2640  
2989 GCCCTCTCCAGGAGACCTGCAAGTTTGTGAGGTGACAGAGCTGAAAGTCAACATCATGTG 3048  
2641 GCCCTCTCCAGGAGACCTGCAAGTTTGTGAGGTGACAGAGCTGAAAGTCAACATCATGTG 2700  
3049 GACACCGCTGAGAGTCACTGACCGCTACCTGTGGATGTGATCCCGTCAACCTGCC 3108

2701 GACACCGCTGAGAGTGCAGTGCAGCGCTACCGTGTGGATGTGATCCCGTCAACCTGCC 2760  
3109 TGGGAGACA GGGGAGAGGCTGCCATCAGCAGAAACA CTTTTCAGAAAGTCAACCGGCT 3168  
2761 TGGGAGACA GGGGAGAGGCTGCCATCAGCAGAAACA CTTTTCAGAAAGTCAACCGGCT 2820  
3169 GTCCCTCTGGGCTCACTTATTTCAAAAGTCTTTGAGTGGAGCATGGGAGGGAGAGCAA 3228  
2821 GTCCCTCTGGGCTCACTTATTTCAAAAGTCTTTGAGTGGAGCATGGGAGGGAGAGCAA 2880  
3229 GCCTCTGACTGTCAACAGACAACCAAACTGGATGTCTCCCACTAA CTTCCAGTTGTCAA 3288  
2881 GCCTCTGACTGTCAACAGACAACCAAACTGGATGTCTCCCACTAA CTTCCAGTTGTCAA 2940  
3289 TGAACCTGATTCTTACTGTCTCTGTGTGATGGAATCTCCACTCGGGCCAGATACAGGATA 3348  
2941 TGAACCTGATTCTTACTGTCTCTGTGTGATGGAATCTCCACTCGGGCCAGATACAGGATA 3000  
3349 CCGACTGACCGTGGGCTTACCCGAG 3408  
3001 CCGACTGACCGTGGGCTTACCCGAG 3060  
3409 TGTCTCAAGTACCACTGAGGAACTCTGACGCTGTGATCTGAGTACACCGTATCTCCCTCGT 3468  
3061 TGTCTCAAGTACCACTGAGGAACTCTGACGCTGTGATCTGAGTACACCGTATCTCCCTCGT 3120  
3469 GGCCTAAAGGGCAACCAAG 3528  
3121 GGCCTAAAGGGCAACCAAG 3180  
3529 TGGGAGCTCTATTTCCACCTTACACACCGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3588  
3181 TGGGAGCTCTATTTCCACCTTACACACCGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
3589 GACCCCTGCTCCAAAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGAGAGAGAGAG 3648  
3241 GACCCCTGCTCCAAAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGAGAGAGAG 3300  
3649 ACCAGAGAGAGTGTGACTCAGACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3708  
3301 ACCAGAGAGAGTGTGACTCAGACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
3709 AGAATACTGTACACCACTCCAAAGTCTGAGAGATGGACAGAGAGAGAGAGAGAGAGAGAGAG 3768  
3361 AGAATACTGTACACCACTCCAAAGTCTGAGAGATGGACAGAGAGAGAGAGAGAGAGAGAGAG 3420  
3769 AAAAAGAGTGTGACACCACTTGTCTCCACCAAACTTGCATCTGAGAGAGAGAGAGAGAGAGAG 3828  
3421 AAAAAGAGTGTGACACCACTTGTCTCCACCAAACTTGCATCTGAGAGAGAGAGAGAGAGAGAG 3480  
3829 CACTGGAGTGTACAGTCTCTGGGAG 3888  
3481 CACTGGAGTGTACAGTCTCTGGGAG 3540  
3889 AATTACCAACACCCCTCAAAAAGGAG 3948  
3541 AATTACCAACACCCCTCAAAAAGGAG 3600  
3949 TGATCAGAGTCTGTGACCTTTGATACCTGAGTCCCGGCTGGAGTACAGTGTCACTGT 4008  
3601 TGATCAGAGTCTGTGACCTTTGATACCTGAGTCCCGGCTGGAGTACAGTGTCACTGT 3660  
4009 TTACACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGATACCAATCATCCAGCTGT 4068  
3661 TTACACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGATACCAATCATCCAGCTGT 3720  
4069 TCCTCTCCCTGACCTGAGTGTGATTCAGCAATGGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4128  
3721 TCCTCTCCCTGACCTGAGTGTGATTCAGCAATGGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
4129 GGCCTCAACCCCTCAATTTGATTTAAACCACTTCTCTGGTGGCTTACTCACCTGTGAAAAA 4188  
3781 GGCCTCAACCCCTCAATTTGATTTAAACCACTTCTCTGGTGGCTTACTCACCTGTGAAAAA 3840



6001 GTATGACACTGGAAATGGTATTGAGTCTCTGCGCACTTCTGTCAGCAACCCAGGTGTGG 6060  
6409 GCAACAAATGATCTTTAGGAACATGGTTTATAGCGGACCAACACCGCCCAACAGGCCAC 6468  
6061 GCAACAAATGATCTTTAGGAACATGGTTTATAGCGGACCAACACCGCCCAACAGGCCAC 6120  
6469 CCCCATAAGGCATAGGCCCAAGACCAATACCGCGGAAATGAGGACCAAGAGCTCTCTCA 6528  
6121 CCCCATAAGGCATAGGCCCAAGACCAATACCGCGGAAATGAGGACCAAGAGCTCTCTCA 6180  
6529 GACAAACATCTCATGGGCCCCATCTCAGGACACTTCTGAGTACATCATTTTCAITGTCATCC 6588  
6181 GACAAACATCTCATGGGCCCCATCTCAGGACACTTCTGAGTACATCATTTTCAITGTCATCC 6240  
6589 TGTGTGACATGATGAAGAACCCCTTACAGTTTCAAGGTTCTGTGGAACCTTTTCAAGTGGCCAC 6648  
6241 TGTGTGACATGATGAAGAACCCCTTACAGTTTCAAGGTTCTGTGGAACCTTTTCAAGTGGCCAC 6300  
6649 TCTGACAGGCTTACAGAGGTCGCCACCTACACATCATATAGTGGAGGCACTGAAGACCA 6708  
6301 TCTGACAGGCTTACAGAGGTCGCCACCTACACATCATATAGTGGAGGCACTGAAGACCA 6360  
6709 GCAGAGCATPAGGTTGGGAAGAGTTGTTACCGTGGGCACTCTGTCAAGAGGCTT 6768  
6361 GCAGAGCATPAGGTTGGGAAGAGTTGTTACCGTGGGCACTCTGTCAAGAGGCTT 6420  
6769 GAACCAACCTACGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTTGCGTGG 6828  
6421 GAACCAACCTACGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTTGCGTGG 6480  
6829 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTTGTGCGAGTGTCTTGGCTT 6888  
6481 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAAACTGTTGTGCGAGTGTCTTGGCTT 6540  
6889 TGGAGTGTCTTATTCAGATGTGATTCATCTAGATGCTGTCATGACATGCTGTGACTA 6948  
6541 TGGAGTGTCTTATTCAGATGTGATTCATCTAGATGCTGTCATGACATGCTGTGACTA 6600  
6949 CAAGATTGGAGAGAGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACTG 7008  
6601 CAAGATTGGAGAGAGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACTG 6660  
7009 TCTTGGGAACGGAAGAGAAATCAAGTGTGACCTTCATGAGGCAACGTTTACATGA 7068  
6661 TCTTGGGAACGGAAGAGAAATCAAGTGTGACCTTCATGAGGCAACGTTTACATGA 6720  
7069 TGGGAGACATACCACTAGGAGAACAGTGGCAGAGGATATCTGGTGCCATTTGCTC 7128  
6721 TGGGAGACATACCACTAGGAGAACAGTGGCAGAGGATATCTGGTGCCATTTGCTC 6780  
7129 CTGCACATGCTTTGGAGGCCAGCGGGCTGGCGTGTGACAACTGGCGCAGACCTGGGG 7188  
6781 CTGCACATGCTTTGGAGGCCAGCGGGCTGGCGTGTGACAACTGGCGCAGACCTGGGG 6840  
7189 TGAACCCAGTCCCAAGGCACTAGTGGCCAGTCTTACAAACAGTATTTCTCAGAGATACCA 7248  
6841 TGAACCCAGTCCCAAGGCACTAGTGGCCAGTCTTACAAACAGTATTTCTCAGAGATACCA 6900  
7249 TCAGAGAACAAACACTAATGTTAATGCCCCTTCAATGAGTGTCTTATGCTTTAGATGTACA 7308  
6901 TCAGAGAACAAACACTAATGTTAATGCCCCTTCAATGAGTGTCTTATGATGTACA 6960  
7309 GGCTGACAGAGAAATTCAGAGAGTAAATCATCTTCCATCCAGAGAACAGCATGT 7368  
6961 GGCTGACAGAGAAATTCAGAGAGTAAATCATCTTCCATCCAGAGAACAGCATGT 7020  
7369 CTCTCTGCAAGATCCATCTTAACTGAGTGTATGAGGAGACCCAGCTTAGAGTTCTTC 7428  
7021 CTCTCTGCAAGATCCATCTTAACTGAGTGTATGAGGAGACCCAGCTTAGAGTTCTTC 7080  
7429 TTTCTTTCTTAAGCCCTTGTCTCTGAGGAAGTGTCTCCAGCTTCACTCAAGCT 7488

Db 7081 TTTCTTTCTTAAGCCCTTGTCTGAGGAAGTCTCCAGCTTCAGCTCAACTCACAGCT 7140  
Qy 7489 TCTCAAGCATCACCTGGGAGTCTCTGAGGGTTTCTCATAAATGAGGCTGCACATT 7548  
Db 7141 TCTCAAGCATCACCTGGGAGTCTCTGAGGGTTTCTCATAAATGAGGCTGCACATT 7200  
Qy 7549 GCCTGTCTGCTTCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7608  
Db 7201 GCCTGTCTGCTTCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7260  
Qy 7609 TTTGTTTGGGATCAATAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7668  
Db 7261 TTTGTTTGGGATCAATAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7320  
Qy 7669 GATATGACCAAAATTTTAACTAGGAGAGTACCACCAACATCTTCTGCTTCACTTAAGTGT 7728  
Db 7321 GATATGACCAAAATTTTAACTAGGAGAGTACCACCAACATCTTCTGCTTCACTTAAGTGT 7380  
Qy 7729 CTGCCCCGCAATCTAGGAAACCAAGCATGATCTTGTCTGATATTTTAAATATCCA 7788  
Db 7381 CTGCCCCGCAATCTAGGAAACCAAGCATGATCTTGTCTGATATTTTAAATATCCA 7440  
Qy 7789 CAGTACT 7795  
Db 7441 CAGTACT 7447

RESULT 6  
AAF21131  
ID AAF21131 standard; DNA; 7680 BP.  
XX  
AC AAF21131;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Human low adenosine antisense oligonucleotide related sequence #2698.  
XX  
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
human; airway disorder; bronchoconstriction; lung inflammation;  
surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
respiratory obstruction; pulmonary obstruction; impeded respiration;  
surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
cancer; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200062736-A2.  
XX  
XX 26-OCT-2000.  
XX  
XX 24-MAR-2000; 2000WO-US008020.  
XX  
XX 06-APR-1999; 99US-0127958P.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX  
XX (NYCE/) NYCE J W.  
XX  
XX Nyce JW;  
XX  
XX WPI; 2000-679539/66.  
XX  
XX Low adenosine (A) content antisense oligonucleotides which do not trigger  
PT adenosine receptors during metabolism, useful e.g. for treating cancers  
PT and respiratory obstructions.  
XX  
XX Disclosure; Page 1026-1028; 1592pp; English.  
XX  
XX The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense



oligonucleotides the A is replaced by a 'Universal' or alternative base.  
(1) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
immunosuppressive, antispasmodic, hypotensive and cytosstatic activities.  
The antisense oligonucleotides and (1) can be used to down-regulate the  
expression and or activity of target polypeptides associated with  
lung/respiratory disorders and malignancies, such as stimulating and  
activating peptide factors and transmitters, transcription factors,  
immunoglobulins and antibodies, antibody receptors, cytokines and  
chemokines, endogenously produced specific and non-specific enzymes,  
binding proteins, adhesion molecules and their receptors, cytokine and  
chemokine receptors, adenosine receptors, bradykinin receptors, central  
nervous system (CNS), and peripheral nervous and non-nervous system  
receptors, CNS and peripheral nervous and non-nervous system peptide  
transmitters, defensins, growth factors, vasoactive peptides and  
receptors, binding proteins and malignancy associated proteins. The  
antisense oligonucleotides may be used in this way to treat disorders  
including respiratory obstruction (especially pulmonary obstruction  
and/or bronchoconstriction) and/or lung inflammation, allergies) and/or  
surfactant hypoproduction which are associated with a disease or  
condition selected from pulmonary vasoconstriction, inflammation,  
allergies, asthma, impeded respiration, respiratory distress syndrome  
(RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
pulmonary transplantation rejection, pulmonary infections, bronchitis,  
and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
fragments and antisense oligonucleotides used in the exemplification of  
the present invention

Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Query Match	95.2%;	Score 7423;	DB 3;	Length 7680;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 7432;	Conservative	0;	Mismatches 15;	Indels 0;
Gaps	0;			
QY	349	GAAGACGACGAGGCGCTCAGCAATGGTTGAGCCCGAGTCCCGGTTGGTGTGACGTCA	408	
DB	1	GAAGACGACGAGGCGCTCAGCAATGGTTGAGCCCGAGTCCCGGTTGGTGTGACGTCA	60	
QY	409	AAGCAAGCCCGGTTGTTATGACATGGAAACACATATCAGATAAATCAACAGTGGGAGCG	468	
DB	61	AAGCAAGCCCGGTTGTTATGACATGGAAACACATATCAGATAAATCAACAGTGGGAGCG	120	
QY	469	GACCTACCTAGGCAAGCGTTGGTTGTTACTGTTATGGAGGAGCCGAGGTTTAACTG	528	
DB	121	GACCTACCTAGGCAAGCGTTGGTTGTTACTGTTATGGAGGAGCCGAGGTTTAACTG	180	
QY	529	CGAACTAAACCTGAGCTGAGAGACTTGTTCACAGTACACTGGGAAACACTTACCG	588	
DB	181	CGAACTAAACCTGAGCTGAGAGACTTGTTCACAGTACACTGGGAAACACTTACCG	240	
QY	589	AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGATCGG	648	
DB	241	AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGATCGG	300	
QY	649	GGCTGGGCGAGGAGAAATAGCTGTACCATCCGAACCCCTGCCATGAAGGGGTCAGTC	708	
DB	301	GGCTGGGCGAGGAGAAATAGCTGTACCATCCGAACCCCTGCCATGAAGGGGTCAGTC	360	
QY	709	CTACAAAGATTGGTGCACACTGGAGGAGACCATGAGACTGGTGGTTACATTTAGAGTG	768	
DB	361	CTACAAAGATTGGTGCACACTGGAGGAGACCATGAGACTGGTGGTTACATTTAGAGTG	420	
QY	769	TGTGTGTTCTGGTAATGGAAAGGAGAAATGGACCTCGAACCCCATAGCTGAGAGTGT	828	
DB	421	TGTGTGTTCTGGTAATGGAAAGGAGAAATGGACCTCGAACCCCATAGCTGAGAGTGT	480	
QY	829	TGATCATGCTGCTGGGACTTCTTATGTGTGCGAGAAACGTTGGGAGAACCCCTACCAAG	888	
DB	481	TGATCATGCTGCTGGGACTTCTTATGTGTGCGAGAAACGTTGGGAGAACCCCTACCAAG	540	
QY	889	CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	948	
DB	541	CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600	

QY	949	TAGAATAGATGCAACAGATCAGGACACAAAGGACATCTCTATAGAAATGGAGACACTTGGAG	1008	
DB	601	TAGAATAGATGCAACAGATCAGGACACAAAGGACATCTCTATAGAAATGGAGACACTTGGAG	660	
QY	1009	CAAGAGGATATCAGGAGAACTCTCCAGTGCATCTGCACAGCAACGGCCGAGGAGA	1068	
DB	661	CAAGAGGATATCAGGAGAACTCTCCAGTGCATCTGCACAGCAACGGCCGAGGAGA	720	
QY	1069	GTGGAAGTGTGAGAGGACACCTCTGTGACAGACACATCGAGCGGATCTGGCCCTTCCAC	1128	
DB	721	GTGGAAGTGTGAGAGGACACCTCTGTGACAGACACATCGAGCGGATCTGGCCCTTCCAC	780	
QY	1129	CGATGTTCTGTCAGCTGTTTACCAACCCGAGCTCACCCGAGCTCTCTCCCTATGGCCA	1188	
DB	781	CGATGTTCTGTCAGCTGTTTACCAACCCGAGCTCACCCGAGCTCTCTCCCTATGGCCA	840	
QY	1189	CTGTGTACAGACAGTGTGTGTCTCTGTGGGATGACAGTGGCTGGAAGACACAGG	1248	
DB	841	CTGTGTACAGACAGTGTGTGTCTCTGTGGGATGACAGTGGCTGGAAGACACAGG	900	
QY	1249	AAATAGCAAAATGCTTTCACCTGCTGGGCAACGAGTTCAGCTGCCAAGAGACAGCTGT	1308	
DB	901	AAATAGCAAAATGCTTTCACCTGCTGGGCAACGAGTTCAGCTGCCAAGAGACAGCTGT	960	
QY	1309	AACCCAGACTTACGGTGGCAACTCAAATGGAGAGCAGTGTCTTACCAATCACCCTACAA	1368	
DB	961	AACCCAGACTTACGGTGGCAACTCAAATGGAGAGCAGTGTCTTACCAATCACCCTACAA	1020	
QY	1369	TGGCAGAGAGTCTCTCTCTGACACACAGAAAGGGGACAGGACGACATCTTTGGTGCAG	1428	
DB	1021	TGGCAGAGAGTCTCTCTCTGACACACAGAAAGGGGACAGGACGACATCTTTGGTGCAG	1080	
QY	1429	CACAACTTCGAATTATGACAGGACCCAGAAATACCTCTTCTGCACAGACCACTGTCTT	1488	
DB	1081	CACAACTTCGAATTATGACAGGACCCAGAAATACCTCTTCTGCACAGACCACTGTCTT	1140	
QY	1489	GGTTCAGACTCGAGGAGGAAATTCGAATGGTCTTGTGCCACTTCCCTTCTTACAA	1548	
DB	1141	GGTTCAGACTCGAGGAGGAAATTCGAATGGTCTTGTGCCACTTCCCTTCTTACAA	1200	
QY	1549	CAACCAATTCACACTGATTGACATTTCTGAGGCGCAGAGAGACAAATGAAGTGTGTGG	1608	
DB	1201	CAACCAATTCACACTGATTGACATTTCTGAGGCGCAGAGAGACAAATGAAGTGTGTGG	1260	
QY	1609	GACACACAGAACTATGATGCGGACAGAAAGTTGGGTTCTGCCCATGGCTGCCACCA	1668	
DB	1261	GACACACAGAACTATGATGCGGACAGAAAGTTGGGTTCTGCCCATGGCTGCCACCA	1320	
QY	1669	GGAAATCTGCACAAACCAATGAAGGGGTCACTGACCGCATTTGGAGATCAGTGGGATAAGCA	1728	
DB	1321	GGAAATCTGCACAAACCAATGAAGGGGTCACTGACCGCATTTGGAGATCAGTGGGATAAGCA	1380	
QY	1729	GCATGACATGGGTCACTATGATGAGTGCAGCTGTGTTGGGAATGGTCTGGGGAATGGAC	1788	
DB	1381	GCATGACATGGGTCACTATGATGAGTGCAGCTGTGTTGGGAATGGTCTGGGGAATGGAC	1440	
QY	1789	ATGCAATGCTTCTGCGAGCTTCAGATCAGTCACTTGTGATGACATCACTTACAAATGT	1848	
DB	1441	ATGCAATGCTTCTGCGAGCTTCAGATCAGTCACTTGTGATGACATCACTTACAAATGT	1500	
QY	1849	GACGACACATTTCCACAAAGCGCTCATGAAGAGGGGCACTGCTGAACTGTACATGCTTCGG	1908	
DB	1501	GACGACACATTTCCACAAAGCGCTCATGAAGAGGGGCACTGCTGAACTGTACATGCTTCGG	1560	
QY	1909	TCAGGTGCGGGGAGGTGGAGGTGATCCCGTCCGACCAATGCCAGGATTCAGAGACTGG	1968	
DB	1561	TCAGGTGCGGGGAGGTGGAGGTGATCCCGTCCGACCAATGCCAGGATTCAGAGACTGG	1620	
QY	1969	GACCTTTTATCAAATGGAGATTCATGGGAGAAAGTATGTGATGGTGTGATGATGATGATG	2028	
DB	1621	GACCTTTTATCAAATGGAGATTCATGGGAGAAAGTATGTGATGGTGTGATGATGATGATG	1680	



Db 3841 TGAGAGAGATGTCGAGAGTGTCAATTTCTCTCTTCCAGCAATGAGTGTCTTAAACAA 3900  
QY TCTCTGCTCGTACAGAAATATGATGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAG 4308  
Db TCTCTGCTCGTACAGAAATATGATGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAG 3960  
QY CACACCTCTTAGAGAGACAGAGAAACAGAGTCTTGTATCCCAACTGGCAATGACATTTTC 4368  
Db CACACCTCTTAGAGAGACAGAGAAACAGAGTCTTGTATCCCAACTGGCAATGACATTTTC 4020  
QY TGATATTAATGCTCAACTCTTTTACTGTGCTACTGATGCTCTCTGAGCCACCATCACTGG 4428  
Db TGATATTAATGCTCAACTCTTTTACTGTGCTACTGATGCTCTCTGAGCCACCATCACTGG 4080  
QY CTACAGATCCGCCATATCCGAGCACTTCAGTGGAGACCTTCGAGAGATCGGGTGCC 4488  
Db CTACAGATCCGCCATATCCGAGCACTTCAGTGGAGACCTTCGAGAGATCGGGTGCC 4140  
QY CCACCTCGGAATCCATCACCTCACCAACCTCACTCCAGGACAGAGTATGTTGTCAG 4548  
Db CCACCTCGGAATCCATCACCTCACCAACCTCACTCCAGGACAGAGTATGTTGTCAG 4200  
QY CATCGTCTCTTATGTCGAGAGGAAAGTCCCTTATGATGTCGCAACAAATCAACAGT 4608  
Db CATCGTCTCTTATGTCGAGAGGAAAGTCCCTTATGATGTCGCAACAAATCAACAGT 4260  
QY TTCGATGTCGAGGACCTGGAAGTGTGCTGCGACCCACACGCTTACTGATCAG 4668  
Db TTCGATGTCGAGGACCTGGAAGTGTGCTGCGACCCACACGCTTACTGATCAG 4320  
QY CTGGATGCTCTGCTGTACAGTGTGATTTACAGGATCACTTACGAGAAACAGGAGG 4728  
Db CTGGATGCTCTGCTGTACAGTGTGATTTACAGGATCACTTACGAGAAACAGGAGG 4380  
QY AAATAGCCCTGTCCAGAGTTCACGTGCTGGGACCAAGTCTACAGTACCATCAGCGG 4788  
Db AAATAGCCCTGTCCAGAGTTCACGTGCTGGGACCAAGTCTACAGTACCATCAGCGG 4440  
QY CCTTAAACCTGGAGTGTATATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4848  
Db CCTTAAACCTGGAGTGTATATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4500  
QY CCCGAGACGACAGCCAAATTTCAATTAATACCGAAGCAAGTGTACAGTACCATCAGCGG 4908  
Db CCCGAGACGACAGCCAAATTTCAATTAATACCGAAGCAAGTGTACAGTACCATCAGCGG 4560  
QY GATGCAAGTGCAGGATGTTTACGAGCAACAGCAATAGTGTCAAGTGGCTGCTTCAAGTTC 4968  
Db GATGCAAGTGCAGGATGTTTACGAGCAACAGCAATAGTGTCAAGTGGCTGCTTCAAGTTC 4620  
QY CCTGTGTTACTGTTACAGAGTAAACCACTCCCAAAATGAGACCAAGACCAACAAAC 5028  
Db CCTGTGTTACTGTTACAGAGTAAACCACTCCCAAAATGAGACCAAGACCAACAAAC 4680  
QY TAAACCTGAGGTCAGATCAACAGAAATGACTATTGAAGCTTTCAGCCCAACAGTGG 5088  
Db TAAACCTGAGGTCAGATCAACAGAAATGACTATTGAAGCTTTCAGCCCAACAGTGG 4740  
QY GTATGTTGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGTCAAGCTTGTGTTGACAG 5148  
Db GTATGTTGTTAGTGTCTATGCTCAGAAATCCAGCGGAGAGTCAAGCTTGTGTTGACAG 4800  
QY TGCAAGTAAACCAATGATGCTGCTTAAAGGACTGGCAATTCATGATGTTGATGTCGATTC 5208  
Db TGCAAGTAAACCAATGATGCTGCTTAAAGGACTGGCAATTCATGATGTTGATGTCGATTC 4860  
QY CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCAGGTTACAGGTTGACCTATCTC 5268  
Db CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCAGGTTACAGGTTGACCTATCTC 4920  
QY GAGCCCTGAGGATGGAATCCATGAGTATTCCTGCACTGTGTTGAGAGACACTGTC 5328  
Db GAGCCCTGAGGATGGAATCCATGAGTATTCCTGCACTGTGTTGAGAGACACTGTC 4980

QY 5329 AGAGCTGCAAGGCTTCAGACCGGGTTCAGATACAGTCAAGTGTGTTGCTTGCACGA 5388  
Db 4981 AGAGCTGCAAGGCTTCAGACCGGGTTCAGATACAGTCAAGTGTGTTGCTTGCACGA 5040  
QY 5389 TGATATGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5448  
Db 5041 TGATATGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5100  
QY 5449 CTGAAGTTCATCAGGTCACACCCCAAGCCTGAGCCCGCAGTGGACACCACTTGT 5508  
Db 5101 CTGAAGTTCATCAGGTCACACCCCAAGCCTGAGCCCGCAGTGGACACCACTTGT 5160  
QY 5509 TCAGTCTCACTGATATCCAGTGCAGTCCACCCCAAGGAGAGCCGACCAATGAAGA 5568  
Db 5161 TCAGTCTCACTGATATCCAGTGCAGTCCACCCCAAGGAGAGCCGACCAATGAAGA 5220  
QY 5569 AATCAACCTTGTCTCTGACAGCTATCCGTGGTGTATCAGGACTTATGTTGGCCACCA 5628  
Db 5221 AATCAACCTTGTCTCTGACAGCTATCCGTGGTGTATCAGGACTTATGTTGGCCACCA 5280  
QY 5629 ATATGAGTGTGCTTATGCTTTAAGGACACTTTTGAACAGCAGACAGCTCAGGAGT 5688  
Db 5281 ATATGAGTGTGCTTATGCTTTAAGGACACTTTTGAACAGCAGACAGCTCAGGAGT 5340  
QY 5689 TGTCACTCTGAGAAATGTGAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCACTCTGAGAAATGTGAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGA 5400  
QY 5749 GACCACTCAGCATTTAGCTGGAGAACCAAGACTGAGCATCTGGGCTTCAAGTTGA 5808  
Db 5401 GACCACTCAGCATTTAGCTGGAGAACCAAGACTGAGCATCTGGGCTTCAAGTTGA 5460  
QY 5809 TGCCGTTCCAGCAATGGCCAGATCCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5868  
Db 5461 TGCCGTTCCAGCAATGGCCAGATCCCAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5520  
QY 5869 CTACACCATCTGCTTACACAGGACACTGACAGATCTACCTGATACCTTTGAA 5928  
Db 5521 CTACACCATCTGCTTACACAGGACACTGACAGATCTACCTGATACCTTTGAA 5580  
QY 5929 TGCAATGCTCGAGCTCCCTGTGTCATCGACGCTCCACTGCCATTTGATGCACTATC 5988  
Db 5581 TGCAATGCTCGAGCTCCCTGTGTCATCGACGCTCCACTGCCATTTGATGCACTATC 5640  
QY 5989 CAACCTGCTTCTGCGCACCAACCAATTCCTTGTGTTGATCATGACGACCGCACG 6048  
Db 5641 CAACCTGCTTCTGCGCACCAACCAATTCCTTGTGTTGATCATGACGACCGCACG 5700  
QY 6049 TGCCAGATTTACCGGCTCATCATCATCAAGTATGAGAGCCCTGGGTCTCTCTCCAGAGAA 6108  
Db 5701 TGCCAGATTTACCGGCTCATCATCATCAAGTATGAGAGCCCTGGGTCTCTCTCCAGAGAA 5760  
QY 6109 GGTCCCTGCGCCCGCTGCTGTCACAGAGTACTATTTACTGGCTGGACCGGAC 6168  
Db 5761 GGTCCCTGCGCCCGCTGCTGTCACAGAGTACTATTTACTGGCTGGACCGGAC 5820  
QY 6169 CGAATATACAAATTTATGTCATTCCTCGAAGATAATCAGAGAGCGAGCCCTGATTTGG 6228  
Db 5821 CGAATATACAAATTTATGTCATTCCTCGAAGATAATCAGAGAGCGAGCCCTGATTTGG 5880  
QY 6229 AAGGAAACAGACAGAGCTTCCCAACTGGTAAACCTTTCACACCCMACTTCTCATGG 6288  
Db 5881 AAGGAAACAGACAGAGCTTCCCAACTGGTAAACCTTTCACACCCMACTTCTCATGG 5940  
QY 6289 ACCAGAGATCTTGGATGTTCTCTTCCAGTTCACCAAGACCCCTTGTGTCACCCACCTTGG 6348  
Db 5941 ACCAGAGATCTTGGATGTTCTCTTCCAGTTCACCAAGACCCCTTGTGTCACCCACCTTGG 6000  
QY 6349 GTATGACACTGGAATGTTATTCAGTTCCTGCACTTCTGTTGTCAGCAACCTGTTGG 6408  
Db 6001 GTATGACACTGGAATGTTATTCAGTTCCTGCACTTCTGTTGTCAGCAACCTGTTGG 6060

6409	GCACAAATGATCTTTTGGAGAACTAGTTTATAGCGGACCAACACCGCCCAACAAACGGCCAC	6469
6061	GCACAAATGATCTTTTGGAGAACTAGTTTATAGCGGACCAACACCGCCCAACAAACGGCCAC	6120
6469	CCCCATAAGGCATAGCCCAAGACCATACCCGCGGAATGTAGCAACAAGAGCTCTCTCTCA	6528
6121	CCCCATAAGGCATAGCCCAAGACCATACCCGCGGAATGTAGCAACAAGAGCTCTCTCTCA	6180
6529	GACAAACATCTCATGGGCCCATTTCCAGACACTCTCTGAGTACATCATTTTCATGTCAATCC	6588
6181	GACAAACATCTCATGGGCCCATTTCCAGACACTCTCTGAGTACATCATTTTCATGTCAATCC	6240
6589	TGTTGGCAGTGTAGAGAACCCCTTACAGTTTCAGGGTTCCTGGAACCTTCTACCAAGTGCAC	6648
6241	TGTTGGCAGTGTAGAGAACCCCTTACAGTTTCAGGGTTCCTGGAACCTTCTACCAAGTGCAC	6300
6649	TCTGACAGGCCCTCACACAGAGTGCACCTACAAACATCATAGTGGAGGCACTGAAGAACA	6708
6301	TCTGACAGGCCCTCACACAGAGTGCACCTACAAACATCATAGTGGAGGCACTGAAGAACA	6360
6709	GCAGAGGCATAAGTTTCGGAGAGAGTTGTACCGTGGGCAACTCTGTCAACGAGAGGCTTT	6768
6361	GCAGAGGCATAAGTTTCGGAGAGAGTTGTACCGTGGGCAACTCTGTCAACGAGAGGCTTT	6420
6769	GAACCAACCTACGGATGACTCGTGCTTTGACCCCTACACAGTGTCCCATTTATGCCGTTGG	6828
6421	GAACCAACCTACGGATGACTCGTGCTTTGACCCCTACACAGTGTCCCATTTATGCCGTTGG	6480
6829	AGATGATGGGAACGAATGTCTGAATCAGGCTTTAACTGTTGTGTCAGTGTCTTAGGCTTT	6888
6481	AGATGATGGGAACGAATGTCTGAATCAGGCTTTAACTGTTGTGTCAGTGTCTTAGGCTTT	6540
6889	TGCAAGTGGTCAATTTACAGATGTCAATCATCTAGATGGTGCATGACAATGGTGTGAACATA	6948
6541	TGCAAGTGGTCAATTTACAGATGTCAATCATCTAGATGGTGCATGACAATGGTGTGAACATA	6600
6949	CAAGATTTGAGAGAAAGTGGGACCGTCAAGTGTGACCTCATGAGGCAACGTTTACCATGA	7008
6601	CAAGATTTGAGAGAAAGTGGGACCGTCAAGTGTGACCTCATGAGGCAACGTTTACCATGA	6660
7009	TCTTTGGGAACGGAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTACCATGA	7068
6661	TCTTTGGGAACGGAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTTTACCATGA	6720
7069	TGGGAAGACA TACCAGTAGGAGAACAGTGGCAGAGGAATATCTCGTGGCCATTTTGCTC	7128
6721	TGGGAAGACA TACCAGTAGGAGAACAGTGGCAGAGGAATATCTCGTGGCCATTTTGCTC	6780
7129	CTGCACATGCTTTGGAGGCCAGCGGGCTCGGCCTGTGACAACTGCCGAGACCTGGGG	7188
6781	CTGCACATGCTTTGGAGGCCAGCGGGCTCGGCCTGTGACAACTGCCGAGACCTGGGG	6840
7189	TGAACCCAGTCCCGAAGGCACACTAGTGGCCAGTCCCTACAACAGTATTTCTCAGAGATACCA	7248
6841	TGAACCCAGTCCCGAAGGCACACTAGTGGCCAGTCCCTACAACAGTATTTCTCAGAGATACCA	6900
7249	TCAGAGAAACAAACACTAATGTATAATTTGCCCAATTTAGTGTCTTATGCTTTAGATGTACA	7308
6901	TCAGAGAAACAAACACTAATGTATAATTTGCCCAATTTAGTGTCTTATGCTTTAGATGTACA	6960
7309	GGCTGACAGAGAGATTTCCGAGAGTAAATCATCTTTTCCAATCCAGAGGAAACAGCATGT	7368
6961	GGCTGACAGAGAGATTTCCGAGAGTAAATCATCTTTTCCAATCCAGAGGAAACAGCATGT	7020
7369	CTCTCTGCCAAGATTCATCTAAACCTGGAGTATGTTTATAGCAGACCCAGCTTTAGAGTTCTTC	7428
7021	CTCTCTGCCAAGATTCATCTAAACCTGGAGTATGTTTATAGCAGACCCAGCTTTAGAGTTCTTC	7080
7429	TTTCTTTCTTAAAGCCCTTTGCTCTGGAGGAGTGTCTCTCAGCTTCAGCTCAACTCACAGCT	7488
7081	TTTCTTTCTTAAAGCCCTTTGCTCTGGAGGAGTGTCTCTCAGCTTCAGCTCAACTCACAGCT	7140
7489	TCTCCAAAGCATACCCCTGGGAGTTTCTGAGGGTTTCTCATAAATGAGGGGTGCAATT	7548

Db	7141	TTCTCCAAAGCATCACCTCGGAGTTTCTCGAGGTTTCTCATAAATGAGGCTGCACATT	7200
Qy	7549	GCCTGTCTTCGTCGGAATPATTCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7608
Db	7201	GCCTGTCTTCGTCGGAATPATTCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA	7260
Qy	7605	TTTGGTTTGGGATCAATAGAAAGCATATGCAGCCAAACCAAGATGCAAAATGTTTTGAAAT	7668
Db	7261	TTTGGTTTGGGATCAATAGAAAGCATATGCAGCCAAACCAAGATGCAAAATGTTTTGAAAT	7320
Qy	7669	GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7728
Db	7321	GATATGACCAAAATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7380
Qy	7729	CTGGCCCGCAATACTGTAGGAAACCAAGCATGATCTTGTACTGTGATATTTTAAATATFCCA	7788
Db	7381	CTGGCCCGCAATACTGTAGGAAACCAAGCATGATCTTGTACTGTGATATTTTAAATATFCCA	7440
Qy	7789	CAGTACT 7795	
Db	7441	CAGTACT 7447	
RESULT 7			
ABAS2689			
ID	ABAS2689	standard; DNA; 7680 BP.	
XX	ABAS2689;		
XX	25-JAN-2002	(first entry)	
XX	Fibronectin	gene SEQ ID NO:75.	
DE			
XX	Human; high bone mass;	HBM gene; Zmax1 gene; chromosome 11; 11c13.3;	
KW	sequence tagged site;	STS; osteoporosis; osteopathic; gene therapy;	
KW	antisense therapy; vaccine;	bone disorder; Paget's disease; sclerostosis;	
KW	osteomalacia; fibrous dysplasia;	ds.	
XX			
OS	Homo sapiens.		
XX			
PN	WO200177327-A1.		
XX			
FD	18-OCT-2001.		
XX			
PF	21-JUN-2000; 2000WO-US016951.		
XX			
PR	05-APR-2000; 2000US-00543771.		
PR	05-APR-2000; 2000US-00544398.		
XX	(GENO-) GENOME THERAPEUTICS CORP.		
PA			
PI	Carulli JP, Little RD, Recker RR, Johnson ML;		
XX			
DR	WPI; 2001-657171/75.		
XX			
PT	New high bone mass (HBM) and Zmax1 genes and proteins useful for		
PT	modulating bone mass for the treatment of e.g. osteoporosis.		
PS	Claim 79; Page 376-378; 443pp; English.		
XX			
CC	The present invention describes the human Zmax1 gene and the high bone		
CC	mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM		
CC	genes have osteopathic activities. The genes can be used in gene therapy,		
CC	antisense therapy and in the production of vaccines. They can be used in		
CC	the diagnosis and treatment of bone disorders including osteoporosis,		
CC	Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.		
CC	ABAS2038 to ABAS2700 and AAG68168 to AAG68193 represent sequences used in		
CC	the exemplification of the present invention		
XX			
SQ	Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;		
	Query Match	95.2%; Score 7423; DB 5; Length 7680;	

Query Match 95.2%; Score 7423; DB 5; Length 7680;

Best Local Similarity 99.8%; Pred. No. 0;		Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	
349	GAAGAGCAAGAGCAGGCTCAGCAAAATGGTTAGCCCCAGTCCCGGTGGCTGTGAGTCA 408		
1	GAAGAGCAAGAGCAGGCTCAGCAAAATGGTTAGCCCCAGTCCCGGTGGCTGTGAGTCA 60		
409	AAGCAAGCCGGTTGTATGACAATGGAAAAACATATCAGATCAATCAACAGTGGAGCG 468		
61	AAGCAAGCCGGTTGTATGACAATGGAAAAACATATCAGATCAATCAACAGTGGAGCG 120		
469	GACCTACCTAGGAATCGTTGGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTG 528		
121	GACCTACCTAGGAATCGTTGGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTG 180		
529	CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGTGAACAAGTACATGGGAAACATTCAG 588		
181	CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGTGAACAAGTACATGGGAAACATTCAG 240		
589	AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGAGTGTACCTGATCGG 648		
241	AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGAGTGTACCTGATCGG 300		
649	GGCTGGCGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTC 708		
301	GGCTGGCGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTC 360		
709	CTACAAGATTGGTGACACTTGGAGAGACACATGAGACTGTGGTTTACATCTAGAGTG 768		
361	CTACAAGATTGGTGACACTTGGAGAGACACATGAGACTGTGGTTTACATCTAGAGTG 420		
769	TGCTGTCTTGGTAATGCAAAAGGAGAAATGGAAGCTGCAAGCCCATAGCTGAGAAGTGT 828		
421	TGCTGTCTTGGTAATGCAAAAGGAGAAATGGAAGCTGCAAGCCCATAGCTGAGAAGTGT 480		
829	TGATCATGCTGCTGGGACTTCTCTATGTGGTGGAGAAAAGCTGGGAGAGCCCTCAAGG 888		
481	TGATCATGCTGCTGGGACTTCTCTATGTGGTGGAGAAAAGCTGGGAGAGCCCTCAAGG 540		
889	CTGATGATGATGATGATGATCTTCCCTGGAGAGGAGCGGACGCATCACTTGCACCTC 948		
541	CTGATGATGATGATGATGATCTTCCCTGGAGAGGAGCGGACGCATCACTTGCACCTC 600		
949	TAGAAATAGATGCAACAGATCAGGACAACAAGACATCTTATGAATTTGGAGACAACCTGGAG 1008		
601	TAGAAATAGATGCAACAGATCAGGACAACAAGACATCTTATGAATTTGGAGACAACCTGGAG 660		
1009	CAAGAGATTAATCGAGGAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 1068		
661	CAAGAGATTAATCGAGGAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 720		
1069	GTGGAAGTGTGAGAGGACACCTCTGTGCAGACCAATCGAGCGGATCTGGCCCTTTCAC 1128		
721	GTGGAAGTGTGAGAGGACACCTCTGTGCAGACCAATCGAGCGGATCTGGCCCTTTCAC 780		
1129	CGATGTTGTGAGCTGTTTACCAACCGAGCTCACCCAGCCCTCTCCCTATGGCCA 1188		
781	CGATGTTGTGAGCTGTTTACCAACCGAGCTCACCCAGCCCTCTCCCTATGGCCA 840		
1189	CTGTGTCAAGACAGTGGTGGTGTACTGTGTGGGATGTCAGTGGTGAAGACACAAGG 1248		
841	CTGTGTCAAGACAGTGGTGGTGTACTGTGTGGGATGTCAGTGGTGAAGACACAAGG 900		
1249	AAATAGCAAAATGCTTTCAGCTGCTGGGCAACGGAGTCACTGCTCCCAAGAGACAGTGT 1308		
901	AAATAGCAAAATGCTTTCAGCTGCTGGGCAACGGAGTCACTGCTCCCAAGAGACAGTGT 960		
1309	AACCCAGACTTACGTTGGCACTCAAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA 1368		
961	AACCCAGACTTACGTTGGCACTCAAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA 1020		
1369	TGGCAGGACGTTTCTACTCTCTGCAACCAAGAGGGGACAGAGCGGACATCTTTGGTGCAG 1428		
1021	TGGCAGGACGTTTCTACTCTCTGCAACCAAGAGGGGCGACAGGACGGACATCTTTGGTGCAG 1080		
1429	CACAACTTCGAATATGAGCAGGACCAAGAAATATCTTTCTTGCAAGACCAACACTGTTTT 1488		
1081	CACAACTTCGAATATGAGCAGGACCAAGAAATATCTTTCTTGCAAGACCAACACTGTTTT 1140		
1489	GGTTCAAGTTCGAGGAGGAAAATTCAAATGGTGGCTTGTGCCACTTCCCTTCTATACAA 1548		
1141	GGTTCAAGTTCGAGGAGGAAAATTCAAATGGTGGCTTGTGCCACTTCCCTTCTATACAA 1200		
1549	CAACCAAAATTAACATGATTCATCTTGAGGGGCAAGAGACAAATGAAGTGGTGTGG 1608		
1201	CAACCAAAATTAACATGATTCATCTTGAGGGGCAAGAGACAAATGAAGTGGTGTGG 1260		
1609	GACCAACAGAACTATGATGCGGACCAAGAAATTTGGTCTTCCGCCATGGCTGCCCAAGA 1668		
1261	GACCAACAGAACTATGATGCGGACCAAGAAATTTGGTCTTCCGCCATGGCTGCCCAAGA 1320		
1669	GGAAATCTGCACAAACCAATGAAGGGGTCAATGATACCGCATTTGAGATACATCAATGAT 1728		
1321	GGAAATCTGCACAAACCAATGAAGGGGTCAATGATACCGCATTTGAGATACATCAATGAT 1380		
1729	GCATGATCGGTTCACATGATGAGGGTGCACGTGTGTTGGAAATGCTGCTGGGSAATGGAC 1788		
1381	GCATGATCGGTTCACATGATGAGGGTGCACGTGTGTTGGAAATGCTGCTGGGSAATGGAC 1440		
1789	ATGCATTTGCCCTACTCTGCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAATG 1848		
1441	ATGCATTTGCCCTACTCTGCAGCTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAATG 1500		
1849	GAACGACATTTCCACAGGCTCATGAAGAGGGGACATGCTGAACTGATCATGCTGCTGG 1908		
1501	GAACGACATTTCCACAGGCTCATGAAGAGGGGACATGCTGAACTGATCATGCTGCTGG 1560		
1909	TCAGGGTCGGGGCAGGTGGAAATGTCATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968		
1561	TCAGGGTCGGGGCAGGTGGAAATGTCATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620		
1969	GACGTTTTATCAAAATTTGGAGATTCATGAGGAGAAATGTCATGCTGATCAGATACCAAG 2028		
1621	GACGTTTTATCAAAATTTGGAGATTCATGAGGAGAAATGTCATGCTGATCAGATACCAAG 1680		
2029	CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAG 2088		
1681	CTACTGCTATGGCGGTGGCAATGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAG 1740		
2089	CTCAAGTGGTCTGTGGAATTTATCACTGAGACTCGAGTCAAGCCCACTCCCAACCC 2148		
1741	CTCAAGTGGTCTGTGGAATTTATCACTGAGACTCGAGTCAAGCCCACTCCCAACCC 1800		
2149	CATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACC 2208		
1801	CATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACC 1860		
2209	TAAATAATCTGTAGGCCGTTTGGAAAGAGCTACCATCAAGGCCACTTAAACTCTTACAC 2268		
1861	TAAATAATCTGTAGGCCGTTTGGAAAGAGCTACCATCAAGGCCACTTAAACTCTTACAC 1920		
2269	CATCAAGGCTGAGGCTGTGTGTATACGAGGGCCAGCTCATCAGATCCAGCAGTA 2328		
1921	CATCAAGGCTGAGGCTGTGTGTATACGAGGGCCAGCTCATCAGATCCAGCAGTA 1980		
2329	CGGCCACCAAGAAATGACTCGCTTTGACTTTCACCAACCAAGCAGCAGCAGTGTGAC 2388		
1981	CGGCCACCAAGAAATGACTCGCTTTGACTTTCACCAACCAAGCAGCAGCAGTGTGAC 2040		
2389	CAGCAACACCGTGAACAGGAGAGCTCCCTTTCTCTCTGTCGCGCACTTCTCAATC 2448		
2041	CAGCAACACCGTGAACAGGAGAGAGCTCCCTTTCTCTCTGTCGCGCACTTCTCAATC 2100		
2449	TGTGACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2508		
2101	TGTGACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2160		



2509	2509	GTCCGGATTTCCGGGTGGGAATATGAGCTGAGTGAGGAGATGAGGCCACAGTACCTGGGA	2568	3599	GACCCCTGCTCCCAAGAAATTTGGTTTTAAGCTGGGTGTATCGACCAAGCCAGGGAGGAGAGGC	3644
2161	Db	GTCCGGATTTCCGGGTGGGAATATGAGCTGAGTGAGGAGGAGATGAGGCCACAGTACCTGGGA	2220	3241	GACCCCTGCTCCCAAGAAATTTGGTTTTAAGCTGGGTGTATCGACCAAGCCAGGGAGGAGAGGC	3300
2569	2569	TCTTCCAAAGCA CAGCCACATCTTGTGTGAACATCCCTGACCTGCTTCTGGCCGAAATATACAT	2528	3649	ACCACGAGAGTGTACTTCAGACTCAGGAACATCGTTGTGTCCGGCTTGCATCTCCAGGAGT	3708
2221	Db	TCTTCCAAAGCA CAGCCACATCTTGTGTGAACATCCCTGACCTGCTTCTGGCCGAAATATACAT	2280	3301	ACCACGAGAGTGTACTTCAGACTCAGGAACATCGTTGTGTCCGGCTTGCATCTCCAGGAGT	3360
2629	2569	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTACTTCCACA	2688	3709	AGAATACGTCTTACACCATCCAAAGTCTTGAGAGATGACAGGAAAAGAGATGCGCAATTTGT	3768
2281	Db	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTACTTCCACA	2340	3361	AGAATACGTCTTACACCATCCAAAGTCTTGAGAGATGACAGGAAAAGAGATGCGCAATTTGT	3420
2689	2569	AAACAACAGCGCTGATGCCCTCTCTGACCCGACCTGTGGACCAAGTTGATGACACCTCAAT	2748	3769	AAAACAAGTGTGTGACACCATTTGCTCCACCAACAAACTTGTGCATCTCGAGGCAAAACCTCTGA	3828
2341	Db	AAACAACAGCGCTGATGCCCTCTCTGACCCGACCTGTGGACCAAGTTGATGACACCTCAAT	2400	3421	AAAACAAGTGTGTGACACCATTTGCTCCACCAACAAACTTGTGCATCTCGAGGCAAAACCTCTGA	3480
2749	2569	TGTTTGTCTGGACAGACCCCGCTCCCATCACACAGGATACAGATAGTCTATTTCGCC	2808	3829	CACCTGGAGTGTCTACAGTCTCTCTGGGAGAGGACACCAACCCAGACATTTACTTGGTTATAG	3888
2401	Db	TGTTTGTCTGGACAGACCCCGCTCCCATCACACAGGATACAGATAGTCTATTTCGCC	2460	3481	CACCTGGAGTGTCTACAGTCTCTCTGGGAGAGGACACCAACCCAGACATTTACTTGGTTATAG	3540
2809	Qy	ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTT	2868	3889	AAATTACACACACCCCTTCAAAACGGCCAGCAGGGAATTTCTTTGGAAGAAGTGTCCATGTC	3948
2461	Db	ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTT	2520	3541	AAATTACACACACCCCTTCAAAACGGCCAGCAGGGAATTTCTTTGGAAGAAGTGTCCATGTC	3600
2869	Qy	CAGTGAATTTGAACTGTGTTCAGTATACATCACTATCTATCTGCTGTGGAAGAAATCA	2928	3949	TGATCAGAGCTCTCTGCACCTTTTGATAAACCTGAGTCCCGGCTCGAGTACAAATGTCAAGTGT	4008
2521	Db	CAGTGAATTTGAACTGTGTTCAGTATACATCACTATCTATCTGCTGTGGAAGAAATCA	2580	3601	TGATCAGAGCTCTCTGCACCTTTTGATAAACCTGAGTCCCGGCTCGAGTACAAATGTCAAGTGT	3660
2929	Qy	AGAAAGTACACCTGTTTTCATTCAACAAGAAACCACTGGCACCCACAGCTCAGATACAGT	2988	4009	TTTACACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT	4068
2581	Db	AGAAAGTACACCTGTTTTCATTCAACAAGAAACCACTGGCACCCACAGCTCAGATACAGT	2640	3661	TTTACACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT	3720
2989	Qy	GGCCTTCCCGAGGACCTGCGATTTGTGGAAGTGA CAGACGTGAAGTCAACCATCTGTG	3048	4069	TCCTCCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTTG	4128
2641	Db	GGCCTTCCCGAGGACCTGCGATTTGTGGAAGTGA CAGACGTGAAGTCAACCATCTGTG	2700	3721	TCCTCCTCCCACTGACCTGCGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTTG	3780
3049	Qy	GACACCGCTGAGGTGACCTGACCGGCTTACCGTGTGATGTGATCCCGTCAACCTGCC	3108	4129	GGCTCCACCCCACTCAATTTGATTAACCAACTTCCTGGTGGTTTACTCACCCTGTGAAAAA	4188
2701	Db	GACACCGCTGAGGTGACCTGACCGGCTTACCGTGTGATGTGATCCCGTCAACCTGCC	2760	3781	GGCTCCACCCCACTCAATTTGATTAACCAACTTCCTGGTGGTTTACTCACCCTGTGAAAAA	3840
3109	Qy	TGGCGAGCAGCGGACGAGCTCCCATCAGCAGGAACACCTTTGCAAGTGCACCGGCT	3168	4189	TGAGGAAGATGTTCCAGAGTTGTCAATTTCTCTTCAGACAATGACAGTGGTCTTTAAACAA	4248
2761	Db	TGGCGAGCAGCGGACGAGCTCCCATCAGCAGGAACACCTTTGCAAGTGCACCGGCT	2820	3841	TGAGGAAGATGTTGAGAGTTGTCAATTTCTCTTCAGACAATGACAGTGGTCTTTAAACAA	3900
3169	Qy	GTCCCTGGGGTCACTTATCTTCAAGTCTTTGAGTGGATGATCCCGTCAACCTGCC	3228	4249	TCCTCTGCTGTGTACAGAAATATGTAGTGTCTCCAGTGTCTACGAAACAAATGAGAG	4308
2821	Db	GTCCCTGGGGTCACTTATCTTCAAGTCTTTGAGTGGATGATCCCGTCAACCTGCC	2880	3901	TCCTCTGCTGTGTACAGAAATATGTAGTGTCTCCAGTGTCTACGAAACAAATGAGAG	3960
3229	Qy	GCCTCTGACTGTCAACAGACAAACAACTGGATGCTCCCACTAACCTTCCAGTTTGTCAA	3288	4309	CHACACTCTTAGAGGAAGACAGAAAAAGAGTCTTTGATTTCCCACTCGCATTTGCTTTTC	4368
2881	Db	GCCTCTGACTGTCAACAGACAAACAACTGGATGCTCCCACTAACCTTCCAGTTTGTCAA	2940	3961	CHACACTCTTAGAGGAAGACAGAAAAAGAGTCTTTGATTTCCCACTCGCATTTGCTTTTC	4020
3289	Qy	TGAAACTGTACTCTGCTGAGATGGACTCGACCTCGGCGCCAGATACAGAGATA	3348	4369	TGATATTACTGCCAACTCTTTTACTGTGCACTGATTTGCTCTCGAGCCACCATCTG	4428
2941	Db	TGAAACTGTACTCTGCTGAGATGGACTCGACCTCGGCGCCAGATACAGAGATA	3000	4021	TGATATTACTGCCAACTCTTTTACTGTGCACTGATTTGCTCTCGAGCCACCATCTG	4080
3349	Qy	CCGACTGACCGTGGGCTTACCCGAGAGAGACAGCCCGAGGAGTACATGTGGGTCCCTC	3408	4429	CTACAGGATTCGCCATCATCCCGAGACATTTCAGTGGGAGACTTCGAGAGATCGGGTGC	4488
3001	Db	CCGACTGACCGTGGGCTTACCCGAGAGAGAGACAGCCCGAGGAGTACATGTGGGTCCCTC	3060	4081	CTACAGGATTCGCCATCATCCCGAGACATTTCAGTGGGAGACTTCGAGAGATCGGGTGC	4140
3409	Qy	TGTTCCAACTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTCGT	3468	4489	CCACTCTCGGAATTCCTATCCCTCACCACCTCACTCCAGGACACAGATATGTGGTCA	4548
3061	Db	TGTTCCAACTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTCGT	3120	4141	CCACTCTCGGAATTCCTATCCCTCACCACCTCACTCCAGGACACAGATATGTGGTCA	4200
3469	Qy	GGCATTAAAGGGCAACAGAGAGCCCAAGCCCTGGAGTCTTTTACCACACTGCGGCC	3528	4549	CATCGTGTCTTTAATGCGAGAGGAAAGTCCCTTTATTTGATTGGCCAAACATCAACAGT	4608
3121	Db	GGCATTAAAGGGCAACAGAGAGCCCAAGCCCTGGAGTCTTTTACCACACTGCGGCC	3180	4201	CATCGTGTCTTTAATGCGAGAGGAAAGTCCCTTTATTTGATTGGCCAAACATCAACAGT	4260
3529	Qy	TGGGAGTCTATTCCACTTACACACCGAGGTGACTG				

Db 4321 CTGGATGCTCTGCTGTCACAGTGAGATATTAACAGATCACTTACGGAGAAACAGGAGG 4380  
 QY 4729 AAATAGCCCTGTCCAGGAGTTTCACTGTGCTGGAGCAAGTCTACAGCTACCACTACGCGG 4788  
 Db 4381 AAATAGCCCTGTCCAGGAGTTTCACTGTGCTGGAGCAAGTCTACAGCTACCACTACGCGG 4440  
 QY 4789 CTTTAAACCTGGAGTTGATATTAACCATCACTGTGTATGCTGTCTACTGGCGGTGGAGACAG 4848  
 Db 4441 CTTTAAACCTGGAGTTGATATTAACCATCACTGTGTATGCTGTCTACTGGCGGTGGAGACAG 4500  
 QY 4849 CCCCAGAGCAGCAGCAATTTCCATTAAATACCGAAACAAATTAACAAACCATCCCA 4908  
 Db 4501 CCCCAGAGCAGCAGCAGCAATTTCCATTAAATTAACCGAAACAAATTAACAAACCATCCCA 4560  
 QY 4909 GATGCAAGTGACCGATGTTTACGAGCAACAGCATTAAGTGTCAAGTGGCTGCTCAAGTTT 4968  
 Db 4561 GATGCAAGTGACCGATGTTTACGAGCAACAGCATTAAGTGTCAAGTGGCTGCTCAAGTTT 4620  
 QY 4969 CCGTGTATGCTGTTACAGAGTAACCAACCACTCCCAAAATTAAGCAGGACCAACAAAAC 5028  
 Db 4621 CCGTGTATGCTGTTACAGAGTAACCAACCACTCCCAAAATTAAGCAGGACCAACAAAAC 4680  
 QY 5029 TAAAACCTGAGCTCCAGATCAAAACAGAAATGACTATTGAAGCTTGCAGCCACAGTGA 5088  
 Db 4681 TAAAACCTGAGCTCCAGATCAAAACAGAAATGACTATTGAAGCTTGCAGCCACAGTGA 4740  
 QY 5089 GTATGTGTTAGTGTCTATGCTCAGATCCAGCCAGGAGAGTCAAGCTCTGGTTCAGAC 5148  
 Db 4741 GTATGTGTTAGTGTCTATGCTCAGATCCAGCCAGGAGAGTCAAGCTCTGGTTCAGAC 4800  
 QY 5149 TGCAGTAACCAACATGATGCGCCCTAAAGGACTGTCATTAAGTGTGATGTCGATTC 5208  
 Db 4801 TGCAGTAACCAACATGATGCGCCCTAAAGGACTGTCATTAAGTGTGATGTCGATTC 4860  
 QY 5209 CATCAAAATGCTTGGAAAGCCACAGGGCAAGTTCCAGTACAGGCTGACCTACTC 5268  
 Db 4861 CATCAAAATGCTTGGAAAGCCACAGGGCAAGTTCCAGTACAGGCTGACCTACTC 4920  
 QY 5269 GAGCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACCTGATGTGAGAGACACTGC 5328  
 Db 4921 GAGCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACCTGATGTGAGAGACACTGC 4980  
 QY 5329 AGAGCTGCAAGGCTCAGACCGGTTCTAGTACACAGTCACTGTGTTGCTTGCAGCA 5388  
 Db 4981 AGAGCTGCAAGGCTCAGACCGGTTCTAGTACACAGTCACTGTGTTGCTTGCAGCA 5040  
 QY 5389 TGATATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGCAACCACTGA 5448  
 Db 5041 TGATATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGCAACCACTGA 5100  
 QY 5449 CTTGAAAGTTCACTCAGGTGACACCCACAGCCCTGAGCCGCTGAGCAACCACTATGT 5508  
 Db 5101 CTTGAAAGTTCACTCAGGTGACACCCACAGCCCTGAGCCGCTGAGCAACCACTATGT 5160  
 QY 5509 TCAGCTCACTGGATATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568  
 Db 5161 TCAGCTCACTGGATATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220  
 QY 5569 AATCAACCTTGTCTCTGACAGCTCACTCGGTGTGTATCAGACTTATGGTGGCCGCCAA 5628  
 Db 5221 AATCAACCTTGTCTCTGACAGCTCACTCGGTGTGTATCAGACTTATGGTGGCCGCCAA 5280  
 QY 5629 ATATGAGTGAAGTGTCTATGCTCTTAAGGACACTTTTGAACGAGACCACTCAGGAGT 5688  
 Db 5281 ATATGAGTGAAGTGTCTATGCTCTTAAGGACACTTTTGAACGAGACCACTCAGGAGT 5340  
 QY 5689 TGTCAACCACTCTGGAGAAATGTCAGCCCAACAAAGAGGCTGCTGTGACAGTGTACTGA 5748  
 Db 5341 TGTCAACCACTCTGGAGAAATGTCAGCCCAACAAAGAGGCTGCTGTGACAGTGTACTGA 5400  
 QY 5749 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCCAAGTTGA 5808

Db 5401 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGA 5460  
 QY 5809 TGCCTGTTCCAGCAATGCGCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCCAGAG 5868  
 Db 5461 TGCCTGTTCCAGCAATGCGCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCCAGAG 5520  
 QY 5869 CTACACCATCACTGCTTACCAACAGGCACTGACTTACCAAGATCTACCTGTACACCTTGA 5928  
 Db 5521 CTACACCATCACTGCTTACCAACAGGCACTGACTTACCAAGATCTACCTGTACACCTTGA 5580  
 QY 5929 TGAACATGCTCGGAGTCTCCCTGTGCTATCAACGCTTCCACTGCAATGATGCAACCTC 5988  
 Db 5581 TGAACATGCTCGGAGTCTCCCTGTGCTATCAACGCTTCCACTGCAATGATGCAACCTC 5640  
 QY 5989 CAACCTGCTTCTTCCGCAACCAACCACTTCTTGTGTGTATCATGCGACGCGCCACG 6048  
 Db 5641 CAACCTGCTTCTTCCGCAACCAACCACTTCTTGTGTGTATCATGCGACGCGCCACG 5700  
 QY 6049 TGCAGGATTAACGGCTATCATCATCAAGTATGAGAGCCTGGTCTCTCCAGAGAACT 6108  
 Db 5701 TGCAGGATTAACGGCTATCATCATCAAGTATGAGAGCCTGGTCTCTCCAGAGAACT 5760  
 QY 6109 GFTCCCTCGCCCGCTGCTGCTACAGAGCTACTATTCTGCGCTGGAACCGGGAAC 6168  
 Db 5761 GFTCCCTCGCCCGCTGCTGCTACAGAGCTACTATTCTGCGCTGGAACCGGGAAC 5820  
 QY 6169 CGAATATACAAATTTATGTCATTGCTGAAAGAAATATCAGAGAGCGGAGCCCTGATTGG 8228  
 Db 5821 CGAATATACAAATTTATGTCATTGCTGAAAGAAATATCAGAGAGCGGAGCCCTGATTGG 5880  
 QY 6229 AAGGAAAGAGACAGAGCTTCCCACTGTAAGCTTCCACACCCCACTTCTCATGG 6288  
 Db 5881 AAGGAAAGAGACAGAGCTTCCCACTGTAAGCTTCCACACCCCACTTCTCATGG 5940  
 QY 6289 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTCTGTCACCCACCTGG 6348  
 Db 5941 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTCTGTCACCCACCTGG 6000  
 QY 6349 GTATGACACTGGAATGTTATTCAGCTTCTGCTGCACTTCTGTCAGCAACCCAGTGTGG 6408  
 Db 6001 GTATGACACTGGAATGTTATTCAGCTTCTGCTGCACTTCTGTCAGCAACCCAGTGTGG 6060  
 QY 6409 GCAACAAATGATCTTTGAGGAACATGTTTGGCGGACCAACCCGCCACCAACGCCAC 6468  
 Db 6061 GCAACAAATGATCTTTGAGGAACATGTTTGGCGGACCAACCCGCCACCAACGCCAC 6120  
 QY 6469 CCCCATAAGCCTAGGCAAGACCATACCCGCGAATGTAGGACAGAGCTCTCTCTCA 6528  
 Db 6121 CCCCATAAGCCTAGGCAAGACCATACCCGCGAATGTAGGACAGAGCTCTCTCTCA 6180  
 QY 6529 GACAACTCTCATGGGCCCTTCCAGGACACTTCTGAGTACATCATTTTCACTGTCATCC 6588  
 Db 6181 GACAACTCTCATGGGCCCTTCCAGGACACTTCTGAGTACATCATTTTCACTGTCATCC 6240  
 QY 6589 TGTGGCACTGATGAAGAACCTTACAGTTCAGGTTCTGGAACCTTCTACCAAGTCCAC 6648  
 Db 6241 TGTGGCACTGATGAAGAACCTTACAGTTCAGGTTCTGGAACCTTCTACCAAGTCCAC 6300  
 QY 6649 TCTGACAGGCTCACCAGAGGTCGCCACTACACATCATATGAGGAGGACTGAAAGACCA 6708  
 Db 6301 TCTGACAGGCTCACCAGAGGTCGCCACTACACATCATATGAGGAGGACTGAAAGACCA 6360  
 QY 6709 GCAGAGGATAGGTTCCGGAAGAGGTTTTCGTTGGGCACTCTGTCAACGAGGCTT 6768  
 Db 6361 GCAGAGGATAGGTTCCGGAAGAGGTTTTCGTTGGGCACTCTGTCAACGAGGCTT 6420  
 QY 6769 GAACCACTACGAGTCACTGCTTGTGACCCCTACACAGTGTCCCATTTATGCGTGG 6828  
 Db 6421 GAACCACTACGAGTCACTGCTTGTGACCCCTACACAGTGTCCCATTTATGCGTGG 6480  
 QY 6829 AGATGAGTGGAAACGAATGTCTGAATCAGGCTTTAACTGTTGTCAGGCTTCTTAGGCTT 6888  
 Db 6481 AGATGAGTGGAAACGAATGTCTGAATCAGGCTTTAACTGTTGTCAGGCTTCTTAGGCTT 6540

y 6889 TGGAGTGGTCAATTTTCAGATGTGATTCATCTAGATGGTCCCATGACATATGGTGTCAACTA 6948  
 b 6541 TGGAGTGGTCAATTTTCAGATGTGATTCATCTAGATGGTCCCATGACATATGGTGTCAACTA 6600  
 y 6949 CAGGATTTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGCTGCACATG 7008  
 b 6601 CAGGATTTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGCTGCACATG 6660  
 y 7009 TCTTGGGACCGGAAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTTACGATGA 7068  
 b 6661 TCTTGGGACCGGAAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTTACGATGA 6720  
 y 7069 TGGGAGACATACCACTGATGAGGAGACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 7128  
 b 6721 TGGGAGACATACCACTGATGAGGAGACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 6780  
 y 7129 CTGCAATCTTTGGAGGCCAGCGGGCTGGCGCTGTGACAACTGCCGACAGCACTGGGGG 7188  
 b 6781 CTGCAATCTTTGGAGGCCAGCGGGCTGGCGCTGTGACAACTGCCGACAGCACTGGGGG 6840  
 y 7189 TGAACCCAGTCCGAGAGGCACTACTGGCCAGTCTCAACACAGATTTCTCAGAGATACA 7248  
 b 6841 TGAACCCAGTCCGAGAGGCACTACTGGCCAGTCTCAACACAGATTTCTCAGAGATACA 6900  
 y 7249 TCAGAGAACAAACACTAAATGTTAAATGCCCAATGAGTGTCTTCATGCTTTCATGATGTACA 7308  
 b 6901 TCAGAGAACAAACACTAAATGTTAAATGCCCAATGAGTGTCTTCATGCTTTCATGATGTACA 6960  
 y 7309 GGTGTACAGAGAGATTCGCGAGAGTAATATCTTTCCATCCAGAGAGGAAACAAGCATGT 7368  
 b 6961 GGTGTACAGAGAGATTCGCGAGAGTAATATCTTTCCATCCAGAGAGGAAACAAGCATGT 7020  
 y 7369 CTCTCTGCCAAGATCCATCTAACTGGAGTGTGTAGCAGACCCAGCTTACAGTTCTTC 7428  
 b 7021 CTCTCTGCCAAGATCCATCTAACTGGAGTGTGTAGCAGACCCAGCTTACAGTTCTTC 7080  
 y 7429 TTTCTTTTAAAGCCCTTTGCTCTGGAGGAAGTCTCCAGCTTACAGTCAACTCAGCT 7488  
 b 7081 TTTCTTTTAAAGCCCTTTGCTCTGGAGGAAGTCTCCAGCTTACAGTCAACTCAGCT 7140  
 y 7489 TCTCCAGATCACTCCCTGGAGTTCTCTGAGGTTTCTCATAAATGAGGGCTGCACATT 7548  
 b 7141 TCTCCAGATCACTCCCTGGAGTTCTCTGAGGTTTCTCATAAATGAGGGCTGCACATT 7200  
 y 7549 GCCTGTTCTGCTTCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTATTTCTAAGA 7608  
 b 7201 GCCTGTTCTGCTTCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTATTTCTAAGA 7260  
 y 7609 TTTGGTTGGATCAATAGGAAGCATATGCAGCAACCAAGATGCAATGTTTGAAT 7668  
 b 7261 TTTGGTTGGATCAATAGGAAGCATATGCAGCAACCAAGATGCAATGTTTGAAT 7320  
 y 7669 GATATGACCAAAATTTAAGTAGGAAAGTCAACCAACCACTTCTGCTTTCACTTAAGTGT 7728  
 b 7321 GATATGACCAAAATTTAAGTAGGAAAGTCAACCAACCACTTCTGCTTTCACTTAAGTGT 7380  
 y 7729 CTGCGCCGCAATCTGTAGGAAACAAGCATGATCTTGTACTGTGATTTTAAATATCCA 7788  
 b 7381 CTGCGCCGCAATCTGTAGGAAACAAGCATGATCTTGTACTGTGATTTTAAATATCCA 7440  
 y 7789 CAGTACT 7795  
 b 7441 CAGTACT 7447

RESULT 8  
 ID ABL67540  
 AC ABL67540 standard; DNA; 7680 BP.  
 CX ABL67540;  
 JT 15-MAY-2002 (first entry)

XX DE Thyroid cancer related gene sequence SEQ ID NO:5877.  
 XX DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 XX DE stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 XX DE cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 XX DE gene; ds.  
 OS Homo sapiens.  
 XX WO200194629-A2.  
 XX 13-DEC-2001.  
 PD 30-MAY-2001; 2001WO-US010838.  
 XX 05-JUN-2000; 2000US-0209473P.  
 XX 05-JUN-2000; 2000US-0209531P.  
 XX 18-SEP-2000; 2000US-0233133P.  
 XX 18-SEP-2000; 2000US-0233617P.  
 XX 20-SEP-2000; 2000US-0234009P.  
 XX 20-SEP-2000; 2000US-0234034P.  
 XX 20-SEP-2000; 2000US-0234052P.  
 XX 22-SEP-2000; 2000US-0234509P.  
 XX 22-SEP-2000; 2000US-0234567P.  
 XX 25-SEP-2000; 2000US-0234923P.  
 XX 25-SEP-2000; 2000US-0234924P.  
 XX 25-SEP-2000; 2000US-0235077P.  
 XX 25-SEP-2000; 2000US-0235082P.  
 XX 25-SEP-2000; 2000US-0235134P.  
 XX 25-SEP-2000; 2000US-0235280P.  
 XX 26-SEP-2000; 2000US-0235637P.  
 XX 26-SEP-2000; 2000US-0235638P.  
 XX 27-SEP-2000; 2000US-0235711P.  
 XX 27-SEP-2000; 2000US-0235720P.  
 XX 27-SEP-2000; 2000US-0235840P.  
 XX 27-SEP-2000; 2000US-0235863P.  
 XX 28-SEP-2000; 2000US-0236028P.  
 XX 28-SEP-2000; 2000US-0236032P.  
 XX 28-SEP-2000; 2000US-0236033P.  
 XX 28-SEP-2000; 2000US-0236034P.  
 XX 28-SEP-2000; 2000US-0236111P.  
 XX 28-SEP-2000; 2000US-0236842P.  
 XX 29-SEP-2000; 2000US-0236891P.  
 XX 02-OCT-2000; 2000US-0237172P.  
 XX 02-OCT-2000; 2000US-0237173P.  
 XX 02-OCT-2000; 2000US-0237278P.  
 XX 02-OCT-2000; 2000US-0237294P.  
 XX 02-OCT-2000; 2000US-0237295P.  
 XX 02-OCT-2000; 2000US-0237316P.  
 XX 03-OCT-2000; 2000US-0237425P.  
 XX 03-OCT-2000; 2000US-0237598P.  
 XX 03-OCT-2000; 2000US-0237604P.  
 XX 03-OCT-2000; 2000US-0237606P.  
 XX 03-OCT-2000; 2000US-0237608P.  
 XX 01-NOV-2000; 2000US-0244867P.  
 XX 01-NOV-2000; 2000US-0245084P.  
 XX (AVAL-) AVALON PHARM.  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX Claim 1; SEQ ID NO 5877; 44pp; English.  
 PS The present invention describes a method (M1) for screening for an anti-  
 CC

neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour

Query Match 95.2%; Score 7423; DB 6; Length 7680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

349 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCA 408  
 1 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCA 60  
 409 AAGCAAGCCCGGTTGTTATGACATGGAACACATATCAGATAATCAACAGTGGGAGCG 468  
 61 AAGCAAGCCCGGTTGTTATGACATGGAACACATATCAGATAATCAACAGTGGGAGCG 120  
 469 GACCTACCTAGGCAATGGTTGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTG 528  
 121 GACCTACCTAGGCAATGGTTGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTG 180  
 529 CGAAGTAAACCTGAGCTGAGAGACTGCTTTGACAGTACACTGGGACACTTACCG 588  
 181 CGAAGTAAACCTGAGCTGAGAGACTGCTTTGACAGTACACTGGGACACTTACCG 240  
 589 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGATCGG 300  
 241 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGATCGG 360  
 649 GGCTGGGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAGGGGTGAGTC 708  
 301 GGCTGGGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCATGAGGGGTGAGTC 360  
 709 CTACAGATTTGGTACACTGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 768  
 361 CTACAGATTTGGTACACTGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 420  
 769 TGTGTCTTTGTTAATGGAAAAGGAGATGAGACTGGTGGTGGTGGTGGTGGTGGTGGTGG 828  
 421 TGTGTCTTTGTTAATGGAAAAGGAGATGAGACTGGTGGTGGTGGTGGTGGTGGTGGTGG 480  
 829 TGATCATGCTGCTGGGACTTCTATGTTGGTGGGAGAAAGCTGGGAGAGCCCTACCAAGG 888  
 481 TGATCATGCTGCTGGGACTTCTATGTTGGTGGGAGAAAGCTGGGAGAGCCCTACCAAGG 540  
 889 CTGGATGATGTTAGTGTACTTGGCTGGGAGAAAGGAGCGGAGCGGAGCGGAGCGGAGCG 948  
 541 CTGGATGATGTTAGTGTACTTGGCTGGGAGAAAGGAGCGGAGCGGAGCGGAGCGGAGCG 600  
 949 TAGAATAGATGCAAGATCAGGACACAGGACATCTTATAGAAATGGAGACACTGGAG 1008  
 601 TAGAATAGATGCAAGATCAGGACACAGGACATCTTATAGAAATGGAGACACTGGAG 660  
 1009 CAAGAGGAGTAATCGAGGAAACCTGCTCGAGTGCATCTGSCACAGCAACCGGCGGAGGAGA 1068  
 661 CAAGAGGAGTAATCGAGGAAACCTGCTCGAGTGCATCTGSCACAGCAACCGGCGGAGGAGA 720  
 1069 GTGGAAGTGTGAGAGGACACCTCTGTGAGACACCATGAGCGGATCTGGCCCTTCCAC 1128

721 GTGGAAGTGTGAGAGGACACACTCTGTGCGACACACATCGAGCGGATCTGSCCCTTCCAC 780  
 1129 CGATTTTCGTGAGAGTGTGTTTACCAACCGAGCTCTACCCCGAGGCTCTCTCTATGGCCA 1188  
 781 CGATTTTCGTGAGAGTGTGTTTACCAACCGAGCTCTACCCCGAGGCTCTCTCTATGGCCA 840  
 1189 CTGTGTACAGACAGTGTGTGTTTACTCTGTGGGATGAGTGGCTGGAAGACACAGG 1248  
 841 CTGTGTACAGACAGTGTGTGTTTACTCTGTGGGATGAGTGGTGTGGAAGACACAGG 900  
 1249 AATTAAGCAAAATGCTTTCAGCTGGGCAACGAGTGCAGTGCCTGCAAGACAGCTGT 1308  
 901 AATTAAGCAAAATGCTTTCAGCTGGGCAACGAGTGCAGTGCCTGCAAGACAGCTGT 960  
 1309 AACCAGACTTACGTTGGCAACTCAATGGAGAGCCATGTCTTACCATTCACCTACAA 1368  
 961 AACCAGACTTACGTTGGCAACTTAAATGGAGAGCCATGTCTTACCATTCACCTACAA 1020  
 1369 TGGCAGGAGTTCCTACTCTGTCACACAGAGGGGCGACAGGACGACATCTTTGGTGCAG 1428  
 1021 TGGCAGGAGTTCCTACTCTGTCACACAGAGGGGCGACAGGACGACATCTTTGGTGCAG 1080  
 1429 CACAACTTCGAATTTATGAGCAGGACCAAGAAATACCTTTCTGCACAGACACACTGTTT 1488  
 1081 CACAACTTCGAATTTATGAGCAGGACCAAGAAATACCTTTCTGCACAGACACACTGTTT 1140  
 1489 GGTTCAGACTTCGAGGAGGAAATTCCAATGGTGGCTTGTGCCACTTCCCTCTCTATACAA 1548  
 1141 GGTTCAGACTTCGAGGAGGAAATTCCAATGGTGGCTTGTGCCACTTCCCTCTCTATACAA 1200  
 1549 CAACCAATTTACCTGATGTCCTTCGAGGCGGAGAGACACATGAGTGTGTGG 1608  
 1201 CAACCAATTTACCTGATGTCCTTCGAGGCGGAGAGACACATGAGTGTGTGG 1260  
 1609 GACCAACACAACTATGATGCGGACACAGAAATTTGGTGTCTGCCCATGGCTGCCCAAG 1668  
 1261 GACCAACACAACTATGATGCGGACACAGAAATTTGGTGTCTGCCCATGGCTGCCCAAG 1320  
 1669 GGAATCTGCACAACTAAGAGGGTCAATGACCGCATTTGGAGATCAGTGGGATAGCA 1728  
 1321 GGAATCTGCACAACTAAGAGGGTCAATGACCGCATTTGGAGATCAGTGGGATAGCA 1380  
 1729 GCATGACATGGGTCACTATGAGGTCACAGTGTGTGGGAATGGTTCGTTGGGAAATGGAC 1788  
 1381 GCATGACATGGGTCACTATGAGGTCACAGTGTGTGGGAATGGTTCGTTGGGAAATGGAC 1440  
 1789 ATGCAATGCTACTGCGAGCTTCAGATCAGTGCATGTTGATGACATCACTTACATGT 1848  
 1441 ATGCAATGCTACTGCGAGCTTCAGATCAGTGCATGTTGATGACATCACTTACATGT 1500  
 1849 GAACGACACATTCACAAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTGG 1908  
 1501 GAACGACACATTCACAAAGCGTCATGAAGAGGGGACATGCTGAACTGTACATGCTTGG 1560  
 1909 TCAGGTCGCGGAGGAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968  
 1561 TCAGGTCGCGGAGGAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620  
 1969 GAGTGTTCATTAATTTGAGATTCATGGGAGAGTATGTCATGTTGTCAGATACCAATGT 2028  
 1621 GAGTGTTCATTAATTTGAGATTCATGGGAGAGTATGTCATGTTGTCAGATACCAATGT 1680  
 2029 CTACTGCTATGGCGGTGGCAATTTGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAA 2088  
 1681 CTACTGCTATGGCGGTGGCAATTTGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAA 1740  
 2089 CTCAAGTGTCTGTGCAAGTATTTTACTGAGACTCCGAGTCCAGCCAACTCCCAACC 2148  
 1741 CTCAAGTGTCTGTGCAAGTATTTTACTGAGACTCCGAGTCCAGCCAACTCCCAACC 1800  
 2149 CATCCAGTGGAAATGCAACCAAGGAGTCTCAATTTCCAAAGTACATTTCTCAGTGGAGAC 2208  
 1801 CATCCAGTGGAAATGCAACCAAGGAGTCTCAATTTCCAAAGTACATTTCTCAGTGGAGAC 1860





Db	4021	TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCTCGAGCCACCACATCACTGG	4080	Db	5101	CCTGAAGTTCACTCAGGTCAACCCACAAAGCCTGAGCGCCAGTGGACACCAACCAATGT	5160
Qy	4429	CTACAGGATCGGCATCATCTCCGAGCACTTCAGTGGAGAGACCTCGAGAAGATCGGGTGCC	4488	Qy	5509	TCAGTCTACTGATATCGAGTGGGGTGACCCCCAAAGAGAACCCGACCAATGAAGA	5568
Db	4081	CTACAGGATCGGCATCATCTCCGAGCACTTCAGTGGAGAGACCTCGAGAAGATCGGGTGCC	4140	Db	5161	TCAGTCTACTGATATCGAGTGGGGTGACCCCCAAAGAGAACCCGACCAATGAAGA	5220
Qy	4489	CCACTCTCGGAATTCATCAACCTCAACCAAGCTCACTCCAGGACAGAGATGTGTCAG	4548	Qy	5569	AATCAACCTTGTCTCTGACAGCTCATCCGTGGTGTATCAGAGCTTATGTGGCCACCA	5628
Db	4141	CCACTCTCGGAATTCATCAACCTCAACCAAGCTCACTCCAGGACAGAGATGTGTCAG	4200	Db	5221	AATCAACCTTGTCTCTGACAGCTCATCCGTGGTGTATCAGAGCTTATGTGGCCACCA	5280
Qy	4549	CATCGTCTCTTAAATGCGAGAGAGAAAGTCCCTTATTTGATGGCCAAACAAATCAACAGT	4608	Qy	5629	ATATGAAGTGAAGTGTCTATGCTTTAAGGACACATTTTGAACAAGCAGACAGCTCAGGAGT	5688
Db	4201	CATCGTCTCTTAAATGCGAGAGAGAAAGTCCCTTATTTGATGGCCAAACAAATCAACAGT	4260	Db	5281	ATATGAAGTGAAGTGTCTATGCTTTAAGGACACATTTTGAACAAGCAGACAGCTCAGGAGT	5340
Qy	4609	TTCTGATGTTCCGAGGACCTTGAAGTTGTTGCTGCGACCCCAACAGCCTACTGATCAG	4668	Qy	5689	TGTCACCACTCTGGAGAAATGTTCAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGA	5748
Db	4261	TTCTGATGTTCCGAGGACCTTGAAGTTGTTGCTGCGACCCCAACAGCCTACTGATCAG	4320	Db	5341	TGTCACCACTCTGGAGAAATGTTCAGCCCAACCAAGAGGGCTCGTGTGACAGATGCTACTGA	5400
Qy	4669	CTGGGATCTCTGCTGTCAAGTGAATATACAGGATCACTTACGGAGAAACAGGAGG	4728	Qy	5749	GACCAACATACCAATAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCCAAGTTGA	5808
Db	4321	CTGGGATCTCTGCTGTCAAGTGAATATACAGGATCACTTACGGAGAAACAGGAGG	4380	Db	5401	GACCAACATACCAATAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCCAAGTTGA	5460
Qy	4729	AAATAGCCCTCTCCAGGAGTTCACTGTCCCTGGAGCAAGTCTACAGTACCATCAGCGG	4788	Qy	5809	TGCGGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGAGTGTGAGA	5868
Db	4381	AAATAGCCCTCTCCAGGAGTTCACTGTCCCTGGAGCAAGTCTACAGTACCATCAGCGG	4440	Db	5461	TGCGGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGAGTGTGAGA	5520
Qy	4789	CCTTAAACCTGGAGTTGATTAACCATCACTGTGTATGCTGCTCACTGGCGTGGAGACAG	4848	Qy	5869	CTACACCACTCACTGGCTTACAAACAGGCACTGAATAAAGATCTTACCTGTACCTTGA	5928
Db	4441	CCTTAAACCTGGAGTTGATTAACCATCACTGTGTATGCTGCTCACTGGCGTGGAGACAG	4500	Db	5521	CTACACCACTCACTGGCTTACAAACAGGCACTGAATAAAGATCTTACCTGTACCTTGA	5580
Qy	4849	CCCCCAAGCAGCAAGCCAAATTTCCATTAATACCGAACAGAAATTCAGAAACCATCCCA	4908	Qy	5929	TGCAATGCTCGAGCTCCCTGTGCTCATCGACGCTCCCACTGCTGCATTTGATGACCAATC	5988
Db	4501	CCCCCAAGCAGCAAGCCAAATTTCCATTAATACCGAACAGAAATTCAGAAACCATCCCA	4560	Db	5581	TGCAATGCTCGAGCTCCCTGTGCTCATCGACGCTCCCACTGCTGCATTTGATGACCAATC	5640
Qy	4909	GATGCAAGTACCGATGTTCAGGACAAACAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC	4968	Qy	5989	CAACCTGCTCTTCTGGCCACCAACCACTTCTTGTGCTGATCATGCGACGCGCCACG	6048
Db	4561	GATGCAAGTACCGATGTTCAGGACAAACAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC	4620	Db	5641	CAACCTGCTCTTCTGGCCACCAACCACTTCTTGTGCTGATCATGCGACGCGCCACG	5700
Qy	4969	CCCTGTTACTGTTACAGAGTAACCAACCACTCCCAAAATGGACAGCAACCAAAAC	5028	Qy	6049	TGCGAGGATACGGGTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGT	6108
Db	4621	CCCTGTTACTGTTACAGAGTAACCAACCACTCCCAAAATGGACAGCAACCAAAAC	4680	Db	5701	TGCGAGGATACGGGTACATCATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGT	5760
Qy	5029	TAAACCTGAGTCCAGATCAACAGAAATGACTATTAAGAGCTTGCAGCCACAGTGA	5088	Qy	6109	GGTCCCTCGGCGCCGCTGGTGTACAGAGGCTTACTATTAAGCTGGCTGGAACCGGGAAC	6168
Db	4681	TAAACCTGAGTCCAGATCAACAGAAATGACTATTAAGAGCTTGCAGCCACAGTGA	4740	Db	5761	GGTCCCTCGGCGCCGCTGGTGTACAGAGGCTTACTATTAAGCTGGCTGGAACCGGGAAC	5820
Qy	5089	GTATGTTGTTAGTGTCTATGCTCAGAAATCCAGAGGAGAGTCAAGCTCTGTTTCAGAC	5148	Qy	6169	CGAATATACAAATTTATGTCATTTGCCCTGAAAGATTAATCAGAGAGGAGCCCTGATGG	6228
Db	4741	GTATGTTGTTAGTGTCTATGCTCAGAAATCCAGAGGAGAGTCAAGCTCTGTTTCAGAC	4800	Db	5821	CGAATATACAAATTTATGTCATTTGCCCTGAAAGATTAATCAGAGAGGAGCCCTGATGG	5880
Qy	5149	TGCAATACCAACATGATGCGCCCTAAAGGACTGGCATTCATGATGTTGATGTCGATTC	5208	Qy	6229	AGGAAAAAGACAGAGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTTCATGG	6288
Db	4801	TGCAATACCAACATGATGCGCCCTAAAGGACTGGCATTCATGATGTTGATGTCGATTC	4860	Db	5881	AGGAAAAAGACAGAGAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTTCATGG	5940
Qy	5209	CATCAAAATTTGCTGGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGGTACCTACTC	5268	Qy	6289	ACCAGAGATCTTGATGTTCTTTCACAGTTCAAAAGACCCCTTTCGTCAACCCCTCG	6348
Db	4861	CATCAAAATTTGCTGGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGGTACCTACTC	4920	Db	5941	ACCAGAGATCTTGATGTTCTTTCACAGTTCAAAAGACCCCTTTCGTCAACCCCTCG	6000
Qy	5269	GAGCCCTGAGGATGGAATCCATGAGTATTCCTGCACTGATGTTGATGTTGATGTTG	5328	Qy	6349	GTATGACCTTGGAAATGTTATTCAGCTTCTGTCACCTTCTGGTCAACCCAGTGTGG	6408
Db	4921	GAGCCCTGAGGATGGAATCCATGAGTATTCCTGCACTGATGTTGATGTTGATGTTG	4980	Db	6001	GTATGACCTTGGAAATGTTATTCAGCTTCTGTCACCTTCTGGTCAACCCAGTGTGG	6060
Qy	5329	AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTTGTTGCTTGCACGA	5388	Qy	6409	GCAACAAATGATCTTTGAGGAAACATGGTTTGGCGGACCAACCGCCCAACAGGCCAC	6468
Db	4981	AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGTTGTTGCTTGCACGA	5040	Db	6061	GCAACAAATGATCTTTGAGGAAACATGGTTTGGCGGACCAACCGCCCAACAGGCCAC	6120
Qy	5389	TGATATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGCAACCACTGA	5448	Qy	6469	CCCCTATAGGCTAGGCGGAGACCTACCCCGGATGTAGGACAGAGAGCTCTCTCTCA	6528
Db	5041	TGATATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGCAACCACTGA	5100	Db	6121	CCCCTATAGGCTAGGCGGAGACCTACCCCGGATGTAGGACAGAGAGCTCTCTCTCA	6180
Qy	5449	CCTGAAGTTCACTCAGTCAACCCCAAGCTGAGCGCCAGTGGACACCAACCAATGT	5508	Qy	6529	GACAAACCATCTCATGGGCCCCCATTTCCAGGACATTTCTGAGTACATCATTTTCTATCC	6588
				Db	6181	GACAAACCATCTCATGGGCCCCCATTTCCAGGACATTTCTGAGTACATCATTTTCTATCC	6240

y	6589	TGTTGGCACTGATGAAGAACCCCTTTACAGTTTCAGGGTTCTGTGAACCTTCTACCAAGTCCAC	6648
b	6241	TGTTGGCACTGATGAAGAACCCCTTACAGTTTCAGGGTTCTGTGAACCTTCTACCAAGTCCAC	6300
y	6649	TCTGACAGCCCTCACAGAGGTGCCACTACAAATCATAGTGGAGGCATCTGAAGACCA	6708
b	6301	TCTGACAGCCCTCACAGAGGTGCCACTACAAATCATAGTGGAGGCATCTGAAGACCA	6360
y	6709	GCAGAGGCATTAAGTTTCGGGAAGAGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTT	6768
b	6361	GCAGAGGCATTAAGTTTCGGGAAGAGTTGTTACCGTGGGCAACTCTGTCAACGAAGGCTT	6420
y	6769	GAACCAACCTACCGGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTAAGCCGTGG	6828
b	6421	GAACCAACCTACCGGATGACTCGTGTCTTGACCCCTACACAGTGTCCCATTAAGCCGTGG	6480
y	6829	AGATGAGTGGGAACGAATGCTTGAATCAGGCTTTAAACTGTTGTGCAAGTGGCTTAGGCTT	6888
b	6481	AGATGAGTGGGAACGAATGCTTGAATCAGGCTTTAAACTGTTGTGCAAGTGGCTTAGGCTT	6540
y	6889	TGGAAATGGTCAATTCAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGACATA	6948
b	6541	TGGAAATGGTCAATTCAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGACATA	6600
y	6949	CAAGATTGGAGAGAGTGGGACCGTCAAGGAGAAAATGGCCAGATGATGAGCTGCACATG	7008
b	6601	CAAGATTGGAGAGAGTGGGACCGTCAAGGAGAAAATGGCCAGATGATGAGCTGCACATG	6660
y	7009	TCTTGGGAACGGAAAAGGAGAAATTCAGTGTGACCTCATGAGGCAACCTGTTCAGATGA	7068
b	6661	TCTTGGGAACGGAAAAGGAGAAATTCAGTGTGACCTCATGAGGCAACCTGTTCAGATGA	6720
y	7069	TGGGAAGACATACCCTTAGGAGAACAGTGGCAGAAAGGAATATCTCGGTGCCATTTGCTC	7128
b	6721	TGGGAAGACATACCCTTAGGAGAACAGTGGCAGAAAGGAATATCTCGGTGCCATTTGCTC	6780
y	7129	CTGCACATGCTTTGGAGGCCAGCGGGCTGTGCAAACTGCGCAGAGACCTGGGG	7188
b	6781	CTGCACATGCTTTGGAGGCCAGCGGGCTGTGCAAACTGCGCAGAGACCTGGGG	6840
y	7189	TGAACCCAGTCCCGAAGGCACACTACTGGCAGTCTTACAAACCAGTATCTCTCAGAGATA	7248
b	6841	TGAACCCAGTCCCGAAGGCACACTACTGGCAGTCTTACAAACCAGTATCTCTCAGAGATA	6900
y	7249	TCAGAGAACAAACATTAATGTTAAATTCGCCAATTTGAGTGTGTTTATGCTTTAGATGTACA	7308
b	6901	TCAGAGAACAAACATTAATGTTAAATTCGCCAATTTGAGTGTGTTTATGCTTTAGATGTACA	6960
y	7309	GGCTGACAGAGAAGATTCCGAGAGTAAATCATCTTTCCAAATCCAGAGAACCAAGCATGT	7368
b	6961	GGCTGACAGAGAAGATTCCGAGAGTAAATCATCTTTCCAAATCCAGAGAACCAAGCATGT	7020
y	7369	CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTAGAGTCTTTC	7428
b	7021	CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTAGAGTCTTTC	7080
y	7429	TTTCTTTTCTTAAGCCCTTTGCTCTGCGAGGAAGTTCTCCAGCTTCAGCTCAACTCAGAGCT	7488
b	7081	TTTCTTTTCTTAAGCCCTTTGCTCTGAGAGGAAGTTCTCCAGCTTCAGCTCAACTCAGAGCT	7140
y	7489	TTCTCCAAAGCATACCCCTGGAGTTTCTGTAGGGTTTTCTCATAAATAGGGCTGCAATT	7548
b	7141	TTCTCCAAAGCATACCCCTGGAGTTTCTGTAGGGTTTTCTCATAAATAGGGCTGCAATT	7200
y	7549	GCCTGTTCTGCTTCGAGATATTCAAATACCGCTCAGTATTTTAAATGAAGTGTATCTTAAGA	7608
b	7201	GCCTGTTCTGCTTCGAGATATTCAAATACCGCTCAGTATTTTAAATGAAGTGTATCTTAAGA	7260
y	7609	TTTTGGTTTGGGATCAATPAGAAAGCATATGCAAGCCCAACCAAGATGCAAAATGTTTTGAAAT	7668
b	7261	TTTTGGTTTGGGATCAATPAGAAAGCATATGCAAGCCCAACCAAGATGCAAAATGTTTTGAAAT	7320

Qy	7669	GATATGACCAAAATTTTAACTAGGAAAGTCACCCAAACACTTCTGCTTTCACCTTAAGTGT	7722
Db	7321	GATATGACCAAAATTTTAACTAGGAAAGTCACCCAAACACTTCTGCTTTCACCTTAAGTGT	7380
Qy	7729	CTGCCCCGCAATACTCTAGGAACAAGCATGATCTTCTGTTACTGTGATATTTTAAATATCCA	7788
Db	7381	CTGCCCCGCAATACTCTAGGAACAAGCATGATCTTCTGTTACTGTGATATTTTAAATATCCA	7440
Qy	7789	CAGTACT 7795	
Db	7441	CAGTACT 7447	

## RESULT 9

ABT11082  
 ID AET11082 standard; cDNA; 7680 BP.  
 XX AC AC  
 XX ABT11082;  
 XX DT 04-DEC-2002 (first entry)  
 XX DE Human breast cancer associated coding sequence SEQ ID NO: 1216.  
 XX KW Human; breast specific gene; breast cancer; differential expression;  
 KW cytostatic; gene therapy; gene; ss.  
 XX OS Homo sapiens.  
 XX WO200259271-A2.  
 XX PD 01-AUG-2002.  
 XX PF 25-JAN-2002; 2002WO-US002176.  
 XX PR 25-JAN-2001; 2001US-0263757P.  
 XX PR 25-APR-2001; 2001US-0286090P.  
 XX PR 23-MAY-2001; 2001US-0292517P.  
 XX (GENE-) GENE LOGIC INC.  
 XX Orr MS, Nation M, Diggins JC, Zeng W;  
 XX WPI; 2002-674803/72.  
 XX Diagnosing breast cancer in a patient comprises detecting the level  
 PT gene expression in cell or tissue samples, where a differential gene  
 PT expression is indicative of breast cancer.  
 XX Claim 1: SEQ ID NO 1216; 260bp + Sequence Listing; English.  
 XX PS

oy



2569 TCTTCAAGCAGCAGCCACTTCTGTGAACATCCTGACCTGCTTCTGCGCCGAAATACAT 2628  
 2221 TCTTCAAGCAGCAGCCACTTCTGTGAACATCCTGACCTGCTTCTGCGCCGAAATACAT 2280  
 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTGTGATCTCTACTCTACA 2688  
 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTGTGATCTCTACTCTACA 2340  
 2689 AACAAACAGCGCCTGATGCGCCCTCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2748  
 2341 AACAAACAGCGCCTGATGCGCCCTCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2400  
 2749 TGTGTCTGCTGAGCAGACCCAGCCTCCATCAACAGGGTACAGATAGTCTATTGCGC 2808  
 2401 TGTGTCTGCTGAGCAGACCCAGCCTCCATCAACAGGGTACAGATAGTCTATTGCGC 2460  
 2809 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAACTCTGAACTCTCGTCAACCT 2868  
 2461 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAACTCTGAACTCTCGTCAACCT 2520  
 2869 CAGTGACTTGCACCTGCTGCTGATTAACATCACTATCTATGCTGTGGAAGAAATCA 2928  
 2521 CAGTGACTTGCACCTGCTGCTGATTAACATCACTATCTATGCTGTGGAAGAAATCA 2580  
 2929 AGAAGTACACCTGCTGCTGATTAACATCAACAGAAACCTGACCCAGCTCAGATACAGT 2988  
 2581 AGAAGTACACCTGCTGCTGATTAACATCAACAGAAACCTGACCCAGCTCAGATACAGT 2640  
 2989 GCCTCTCCAGGAGCCTGCGAGTTGTGGAAGTGCAGACAGTGAAGGTCAACCATCATGTG 3048  
 2641 GCCTCTCCAGGAGCCTGCGAGTTGTGGAAGTGCAGACAGTGAAGGTCAACCATCATGTG 2700  
 3049 GACACCGCTGAGAGTGCAGTGCAGCCGCTACCGTGTGATGTGATCCCGCTCAACCTGCC 3108  
 2701 GACACCGCTGAGAGTGCAGTGCAGCCGCTACCGTGTGATGTGATCCCGCTCAACCTGCC 2760  
 3109 TGCGAGACCGGACAGAGCTGCCATCAGCAGAAACACTTTGACAGAGTCAACCGGCT 3168  
 2761 TGCGAGACCGGACAGAGCTGCCATCAGCAGAAACACTTTGACAGAGTCAACCGGCT 2820  
 3169 GTCCCTCGGGTCACTTATTAATTAAGTCTTTGAGTGGAGTGGAGGAGGAGCAA 3228  
 2821 GTCCCTCGGGTCACTTATTAAGTCTTTGAGTGGAGTGGAGGAGGAGCAA 2880  
 3229 GCCTCTGACTGCTCAACAGACAAACCTGAGTGTCTCCACTTAACCTCCAGTTGTCAA 3288  
 2881 GCCTCTGACTGCTCAACAGACAAACCTGAGTGTCTCCACTTAACCTCCAGTTGTCAA 2940  
 3289 TGAACCTGATTTACTGCTGCTGAGTGGAGTCCACTCGGCGCCAGATTAACAGGATA 3348  
 2941 TGAACCTGATTTACTGCTGCTGAGTGGAGTCCACTCGGCGCCAGATTAACAGGATA 3000  
 3349 CCGACTGACCGTGGGCTTACCCGAGAGGACAGCCAGCAGTACAAATGGGCTCCCTC 3408  
 3001 CCGACTGACCGTGGGCTTACCCGAGAGGACAGCCAGCAGTACAAATGGGCTCCCTC 3060  
 3409 TGTCTCAAGTACCCCTGAGGAATCTGACGCTGCTGAGTACACCGTATCCCTGCT 3468  
 3061 TGTCTCAAGTACCCCTGAGGAATCTGACGCTGCTGAGTACACCGTATCCCTGCT 3120  
 3469 GGCCATTAAGGGCAACCAAGAGGCGCCAAAGCCACTGGAGTCTTTTACCACCTCGAGC 3528  
 3121 GGCCATTAAGGGCAACCAAGAGGCGCCAAAGCCACTGGAGTCTTTTACCACCTCGAGC 3180  
 3529 TGGAGCTCTATTCCACTTACCAACCGAGGCTGACTGAGACCAACCATTTGTATCACATG 3588  
 3181 TGGAGCTCTATTCCACTTACCAACCGAGGCTGACTGAGACCAACCATTTGTATCACATG 3240  
 3589 GACGCTGCTCCAAAGAAATGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGC 3648  
 3241 GACGCTGCTCCAAAGAAATGGTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGC 3300

QY 3649 ACCACGAGAGTGACTTCACTCAGAGCACTCGTGTGCTCGGCTTGACTCAGAGT 3708  
 DB 3301 ACCACGAGAGTGACTTCACTCAGAGCACTCGTGTGCTCGGCTTGACTCAGAGT 3360  
 QY 3709 AGATACGTCTACACCAATCCAAAGTCTGAGAGTGGACAGAAAGAGATCGGCAATGT 3768  
 DB 3361 AGATACGTCTACACCAATCCAAAGTCTGAGAGTGGACAGAAAGAGATCGGCAATGT 3420  
 QY 3769 AAAACAAAGTGGTGAACACCATTTGCTCCACCAACAACTTGCATCTGGAGCAAAACCTGA 3828  
 DB 3421 AAAACAAAGTGGTGAACACCATTTGCTCCACCAACAACTTGCATCTGGAGCAAAACCTGA 3480  
 QY 3829 CACTGGAGTGTCAAGTCTCTGGAGAGGAGCAACACCCAGACATTAATGCTGTATAG 3888  
 DB 3481 CACTGGAGTGTCAAGTCTCTGGAGAGGAGCAACACCCAGACATTAATGCTGTATAG 3540  
 QY 3889 AATTACCAACACCCCTACAAACCGGAGGAGGAAATTTCTTGGAGAAAGTGTCTCATGC 3948  
 DB 3541 AATTACCAACACCCCTACAAACCGGAGGAGGAAATTTCTTGGAGAAAGTGTCTCATGC 3600  
 QY 3949 TGATCAGAGTCTCTGACATTTTGTATTAACCTGAGTCCCGGCTGGAGTACATGTCAGTGT 4008  
 DB 3601 TGATCAGAGTCTCTGACATTTTGTATTAACCTGAGTCCCGGCTGGAGTACATGTCAGTGT 3660  
 QY 4009 TTACACTGTCAAGGATGACAAAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068  
 DB 3661 TTACACTGTCAAGGATGACAAAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720  
 QY 4069 TCCTCTCCACTGACCTGCGATTCACCAACATTTGTCAGACACCAATGCTGTCACTG 4128  
 DB 3721 TCCTCTCCACTGACCTGCGATTCACCAACATTTGTCAGACACCAATGCTGTCACTG 3780  
 QY 4129 GGCTCCACCCCTCCTCAATTTGATTTAAACAACTTCTGCTGCTTACTCACTGTGAAAAA 4188  
 DB 3781 GGCTCCACCCCTCCTCAATTTGATTTAAACAACTTCTGCTGCTTACTCACTGTGAAAAA 3840  
 QY 4189 TGAGGAAGTGTGACAGATGTCATTTCTCTTACAGAAATGCTGAGTGTCTTTAAACAA 4248  
 DB 3841 TGAGGAAGTGTGACAGATGTCATTTCTCTTACAGAAATGCTGAGTGTCTTTAAACAA 3900  
 QY 4249 TCTCCTGCTGTCAGACATATGATGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAG 4308  
 DB 3901 TCTCCTGCTGTCAGACATATGATGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAG 3960  
 QY 4309 CACACCTCTTAGAGGAAGACAGAAACAGTCTTGATTTCCCAACTGGCATTTGCTTTTC 4368  
 DB 3961 CACACCTCTTAGAGGAAGACAGAAACAGTCTTGATTTCCCAACTGGCATTTGCTTTTC 4020  
 QY 4369 TGATATTACTGCGAAGTCTTTTACTGTGCTGATTTGCTTCTCGAGCCACCATCTG 4428  
 DB 4021 TGATATTACTGCGAAGTCTTTTACTGTGCTGATTTGCTTCTCGAGCCACCATCTG 4080  
 QY 4429 CTACAGGATCGCCATCATCCCGAGCACTTCAGTGGAGACCTCGAGAAAGATCGGTTGCC 4488  
 DB 4081 CTACAGGATCGCCATCATCCCGAGCACTTCAGTGGAGACCTCGAGAAAGATCGGTTGCC 4140  
 QY 4489 CCACCTCTCGGAATTCATCACCCTCAACACCTTCTCCAGGACACAGATGTTGTCAG 4548  
 DB 4141 CCACCTCTCGGAATTCATCACCCTCAACACCTTCTCCAGGACACAGATGTTGTCAG 4200  
 QY 4549 CATCGTGTCTTAAATGGCAGAGGAAAGTCTTATTTGATTTGGCCCAACAACTCAACAGT 4608  
 DB 4201 CATCGTGTCTTAAATGGCAGAGGAAAGTCTTATTTGATTTGGCCCAACAACTCAACAGT 4260  
 QY 4609 TTCTGATGTTCCGAGGACCTGGAAGTGTGTCGACACCCCAACAGCTTACTGATCAG 4668  
 DB 4261 TTCTGATGTTCCGAGGACCTGGAAGTGTGTCGACACCCCAACAGCTTACTGATCAG 4320  
 QY 4669 CTGGAGTGTCTCTCTGTCAAGTGTGATATTACAGGATCACTTACGAGAAACAGAGG 4728  
 DB 4321 CTGGAGTGTCTCTCTGTCAAGTGTGATATTACAGGATCACTTACGAGAAACAGAGG 4380  
 QY 4729 AAATAGCCCTGTCCAGGAGTTTCACTGTGCTGGAGCAAGTCTTACAGCTTACCTACGCGG 4788

5461	TGCGTTTCAGCCAAATGGCCGACGCTCCAAATCCAGAGAACCATCAAGCCGACGATGTCAGAAG	5520
5869	CTACACCAATCACTGGCTTTACAAACGAGGCACTGACATPACAAGATCTACTCTGTACACCTTGAA	5928
5521	CTACACCAATCAACAGGTTTACAAACGAGCACTGACATPACAAGATCTACTCTGTACACCTTGAA	5580
5929	TGACAAATGCTCGGAGCTCCCTCTGTGGTCAATGAGCGCTCCACTGCGCAATGATGACACCATC	5988
5591	TGACAAATGCTCGAGCTCCCTCTGTGGTCAATGAGCGCTCCACTGCGCAATGATGACACCATC	5640
5989	CAACCTCGGTTTCTCGGCAACACACACCAATTCCTTTGCTGTGTATCATGTGCGACGCGCCACG	6048
5641	CAACCTCGGTTTCTCGGCAACACACACCAATTCCTTTGCTGTGTATCATGTGCGACGCGCCACG	5700
6049	TGCCAGAGATTACCGGCTACATCATCAAGATATGAGAAGCCTCGGCTCTCTCTCCACGAGAAGT	6108
5701	TGCCAGAGATTACCGGCTACATCATCAAGATATGAGAAGCCTCGGCTCTCTCTCCACGAGAAGT	5760
6109	GGTCCCTCGGCCCCGCCCTTGGTGTACAGAGGCTACTATTACTTGGCCCTGBNACCGGGAAC	6168
5761	GGTCCCTCGGCCCCGCCCTTGGTGTACAGAGGCTACTATTACTTGGCCCTGBNACCGGGAAC	5820
6169	CGAATATACAATTTATGTCTATTGGCCCTGAAGAATTAATCAGAAGAGCGAGCCCTGATTGG	6228
5821	CGAATATACAATTTATGTCTATTGGCCCTGAAGAATTAATCAGAAGAGCGAGCCCTGATTGG	5880
6229	AAGGAAAAGAGACAGAGGCTTCCCAACTGTGTACCCCTTCCACACCCCAATCTTCATGG	6288
5881	AAGGAAAAGAGACAGAGGCTTCCCAACTGTGTACCCCTTCCACACCCCAATCTTCATGG	5940
6289	ACCAGAGATCTTGGATGTTCCCTTCCACAGTTCAAAAGAGCCCTTTGTCACCCACCCCTGG	6348
5941	ACCAGAGATCTTGGATGTTCCCTTCCACAGTTCAAAAGAGCCCTTTGTCACCCACCCCTGG	6000
6349	GTATGACACTGGAAAATGATTTACGCTTCCCTGGCACTTCTGGTCAGCAACCCAGTGTTGG	6408
6001	GTATGACACTGGAAAATGATTTACGCTTCCCTGGCACTTCTGGTCAGCAACCCAGTGTTGG	6060
6409	GCAACAAATGATCTTTGAGGAAACATGGTTTTAGGCGGACCAACCGCCCAACAAAGCGCAC	6468
6061	GCAACAAATGATCTTTGAGGAAACATGGTTTTAGGCGGACCAACCGCCCAACAAAGCGCAC	6120
6469	CCCATAGGCATAGCCACAGACCATACCGCGCAATGTAGGACAAGAAGCTCTCTCTCA	6528
6121	CCCATAGGCATAGCCACAGACCATACCGCGCAATGTAGGACAAGAAGCTCTCTCTCA	6180
6529	GACACCAATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCTTCATGTCAATCC	6588
6181	GACACCAATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCTTCATGTCAATCC	6240
6589	TGTTGGCACTGATGAGAACCCCTTACAGTTACAGGTTTCTCGAACTTCTACCAAGTGCCAC	6648
6241	TGTTGGCACTGATGAGAACCCCTTACAGTTACAGGTTTCTCGAACTTCTACCAAGTGCCAC	6300
6649	TCGTGACAGGCTCACAGAGGTCACATACATCATAGTGTGAGGACCACTGAAAGACCA	6708
6301	TCGTGACAGGCTCACAGAGGTCACATACATCATAGTGTGAGGACCACTGAAAGACCA	6360
6709	CGAGAGGCATAAGGTTTCSGGAAGAGGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT	6768
6361	CGAGAGGCATAAGGTTTCSGGAAGAGGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT	6420
6769	GAAACCAACTTACGAGTACCTGTGCTTTGACCCCTTACAGTGTCCCATATGCGCGTTGG	6828
6421	GAAACCAACTTACGAGTACCTGTGCTTTGACCCCTTACAGTGTCCCATATGCGCGTTGG	6480
6829	ACATCAGTGGGAAACGAATGTCGTGAATCAGGCTTTAAACTGTTGTGCCAGTGTCTAGGCTT	6888
6481	ACATCAGTGGGAAACGAATGTCGTGAATCAGGCTTTAAACTGTTGTGCCAGTGTCTAGGCTT	6540
6889	TGGAAAGTGGTCAATTCAGATGTGATTTCACTCAGATGGTGCCCATGCAATGGTGTGAACTA	6948
6541	TGGAAAGTGGTCAATTCAGATGTGATTTCACTCAGATGGTGCCCATGCAATGGTGTGAACTA	6600



6949 CAAGATTGGAGAGAGTGGAGACCGTCCAGGAGAGAAATGGCCAGATGATGAGTGCACATG 7008  
 6601 CAAGATTGGAGAGAGTGGAGACCGTCCAGGAGAGAAATGGCCAGATGATGAGTGCACATG 6660  
 7009 TCTTGGGAACGGAAAGAGAAATCAAGTGTGACCCCTCATGAGCAACGTTGTTACGATGA 7068  
 6661 TCTTGGGAACGGAAAGAGAAATCAAGTGTGACCCCTCATGAGCAACGTTGTTACGATGA 6720  
 7069 TGGGAAGACATACACGCTAGGAGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTC 7128  
 6721 TGGGAAGACATACACGCTAGGAGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTC 6780  
 7129 CTGCACATGCTTTCGAGCGCAGCGGCTGCGCTGTGACAACTGCGCGAGACCTGGGG 7188  
 6781 CTGCACATGCTTTCGAGCGCAGCGGCTGCGCTGTGACAACTGCGCGAGACCTGGGG 6840  
 7189 TGAACCCAGTCCGAGGCACTACTGSCCACTCTTCAACACAGTATTCTCAGAGATACCA 7248  
 6841 TGAACCCAGTCCGAGGCACTACTGSCCACTCTTCAACACAGTATTCTCAGAGATACCA 6900  
 7249 TCAGAGAACAACTAATGTTAATGCCCAATGAGTGTCTCATGCTTTCAGATGTACA 7308  
 6901 TCAGAGAACAACTAATGTTAATGCCCAATGAGTGTCTCATGCTTTCAGATGTACA 6960  
 7309 GGCTGACAGAGAGATCCCGAGAGTAAATCATCTTTCCAACTCCAGAGGAACCAAGCATGT 7368  
 6961 GGCTGACAGAGAGATCCCGAGAGTAAATCATCTTTCCAACTCCAGAGGAACCAAGCATGT 7020  
 7369 CTCTCTCCAGATCCATCTTAACTGAGAGTATGTTAGCAGACCCAGCTTAGAGTTCTTC 7428  
 7021 CTCTCTCCAGATCCATCTTAACTGAGAGTATGTTAGCAGACCCAGCTTAGAGTTCTTC 7080  
 7429 TTTCTTTCTTAAAGCCCTTGTCTCGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAGCT 7488  
 7081 TTTCTTTCTTAAAGCCCTTGTCTCGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAGCT 7140  
 7489 TCTCCAGATCACCTGGAGTTCTCGAGGTTCTCGAGGTTCTCATAAATGAGGGCTGCACATT 7548  
 7141 TCTCCAGATCACCTGGAGTTCTCGAGGTTCTCGAGGTTCTCATAAATGAGGGCTGCACATT 7200  
 7549 GCGTGTCTGCTTCCAGATTTCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA 7608  
 7201 GCGTGTCTGCTTCCAGATTTCAATACCGCTCAGTATTTTAAATGAAGTGATCTTAAGA 7260  
 7609 TTTGGTTTGGATCAATAGGAAGCATATGCGCAACCAAGATGCAAAATGTTTGAAT 7668  
 7261 TTTGGTTTGGATCAATAGGAAGCATATGCGCAACCAAGATGCAAAATGTTTGAAT 7320  
 7669 GATATGACCAAAATTTTAACTAGGAAGTCCACCAACACTTCTGCTTTCACCTAAGTGT 7728  
 7321 GATATGACCAAAATTTTAACTAGGAAGTCCACCAACACTTCTGCTTTCACCTAAGTGT 7380  
 7729 CTGCGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7788  
 7381 CTGCGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7440  
 7789 CAGTACT 7795  
 7441 CAGTACT 7447

RESULT 10  
 AB295825  
 ID AB295825 standard; DNA; 7680 BP.

XX AB295825;

XX AB295825;

XX 17-OCT-2003 (first entry)

DT Human nucleic acid sequence.

XX Human; antisense; lung dysfunction; nasal airway dysfunction;

XX Human; antisense; lung dysfunction; nasal airway dysfunction;

KW Human; antisense; lung dysfunction; nasal airway dysfunction;

antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;  
 antiallergic; hypotensive; immunosuppressive; cytostatic; gene therapy;  
 antisense gene therapy; respiratory; lung; adenine sensitivity;  
 adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;  
 lung inflammation; respiratory disease; ds.

OS Homo sapiens.

XX WO200285308-A2.

XX 31-OCT-2002.

XX 23-APR-2002; 2002WO-US013135.

XX 24-APR-2001; 2001US-0286137P.

XX (EPIG-) EPIGENESIS PHARM INC.

XX Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;

XX Miller S, Tang L, Shahabuddin S;

XX WPI; 2003-229219/22.

Pharmaceutical composition for treating ailments associated with impaired  
 respiration, has oligo(s) antisense to specific gene(s) or its  
 corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or  
 ubiquinone.

XX Disclosure, SEQ ID NO 12067; 872pp; English.

The invention relates to a novel pharmaceutical composition, which has a  
 first active agent comprising an oligonucleotide antisense to the  
 initiation codon, coding region, 5' or 3' end genomic flanking regions,  
 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of  
 junctions of genes encoding a polypeptide associated with lung and/or  
 nasal airway dysfunction and a second active agent comprising an  
 antiinflammatory steroid and ubiquinone. A composition of the invention  
 has antiinflammatory, antiallergic, antiallergic, hypotensive,  
 immunosuppressive, and cytostatic activity. The composition may have a  
 use in antisense gene therapy. The composition is useful for treating or  
 preventing a respiratory, lung or malignant disease or condition, also  
 for enhancing the prophylactic or therapeutic respiratory effect of an  
 antiinflammatory steroid in a subject, for reducing or depleting levels  
 of, or reducing sensitivity to adenosine, reducing levels of ubiquinone or  
 receptor, producing bronchodilation, increasing levels of ubiquinone or  
 lung surfactant in a subject's tissue, or treating bronchoconstriction,  
 lung inflammation, lung allergies, or a respiratory disease or condition.  
 Note: The sequence data for this patent is not represented in the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Query Match 95.2%; Score 7423; DB 7; Length 7680;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 349 GAAGAGCAGAGGAGGCTCAGCAATGTTTACGCCCCCAGTCCCCGGTGGTGTGTCAGTCA 408

DB 1 GAAGAGCAGAGGAGGCTCAGCAATGTTTACGCCCCCAGTCCCCGGTGGTGTGTCAGTCA 60

QY 409 AAGCAAGCCCGGTTCTTATGCAATGGAACACACTATCAGATAAATCAACAGTGGAGCG 468

DB 61 AAGCAAGCCCGGTTCTTATGCAATGGAACACACTATCAGATAAATCAACAGTGGAGCG 120

QY 469 GACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGAGGAGAGCCGAGGTTTAACTG 528

DB 121 GACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGAGGAGAGCCGAGGTTTAACTG 180

QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGCAACACTTACCG 588

DB 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGCAACACTTACCG 240

589	AGTGGGTGACATTATGAGCGTCTAAAGACCTCCATGATCTGGGACTGTACTCGATCGG	648
241	AGTGGGTGACATTATGAGCGTCTAAAGACCTCCATGATCTGGGACTGTACTCGATCGG	300
649	GSGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAAACCGCTGCCATGAAGGGGGTCACTC	708
301	GSGCTGGGCGAGGAGAAATHAGCTGTATCCATCGCAAACCGCTGCCATGAAGGGGGTCACTC	360
709	CTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTGTTTACATGTTTAGAGTG	768
361	CTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTGTTTACATGTTTAGAGTG	420
769	TGTGTGCTTGTGTATGGAAGAAAGGAGAAATGGACCTTCGCAAGCCCATAGCTGAGAAAGTGT	828
421	TGTGTGCTTGTGTATGGAAGAAAGGAGAAATGGACCTTCGCAAGCCCATAGCTGAGAAAGTGT	480
829	TGATCATGCTGTGGGACATTCCTATGTGTGCGAGAAACGTGGGAGAAAGCCCTACCAAGG	888
481	TGATCATGCTGTGGGACATTCCTATGTGTGCGAGAAACGTGGGAGAAAGCCCTACCAAGG	540
889	CTGGATGATGTTAGATTGTACTTTGCTGGGAGNAGGCAGCGGACGCATCATCTTGCACTTC	948
541	CTGGATGATGTTAGATTGTACTTTGCTGGGAGNAGGCAGCGGACGCATCATCTTGCACTTC	600
949	TAGAAATAGATGCAACCGATCAGGACACAAGAGCATCTTATAGAAATGGAGACACCTGGAG	1008
601	TAGAAATAGATGCAACCGATCAGGACACAAGAGCATCTTATAGAAATGGAGACACCTGGAG	660
1009	CAAGAAAGGATTAATCGAGGAAAACCTGCTCAGTGCATCTGCCAGAGCAAGCGCGAGAGAGA	1068
661	CAAGAAAGGATTAATCGAGGAAAACCTGCTCAGTGCATCTGCCAGAGCAAGCGCGAGAGAGA	720
1069	GTGGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCGAGCGGATCTGGCCCTTCAC	1128
721	GTGGAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCGAGCGGATCTGGCCCTTCAC	780
1129	CGATGTTGCTGACAGTGTGTTTACAAACCGAGCGCTCACCCCGAGCCTCTCCCTATGGCCA	1188
781	CGATGTTGCTGACAGTGTGTTTACAAACCGAGCGCTCACCCCGAGCCTCTCCCTATGGCCA	840
1189	CTGTGTCCACAGACAGTGTGTGCTACTCTGTGGGGATGTCAGTGGCTCAAGACACAAGG	1248
841	CTGTGTCCACAGACAGTGTGTGCTACTCTGTGGGGATGTCAGTGGCTCAAGACACAAGG	900
1249	AAATAAGCAAAATCTTTGCAACGTCCTGGGCAACCGAGTCAGCTGCCAAGACAGCTGT	1308
901	AAATAAGCAAAATCTTTGCAACGTCCTGGGCAACCGAGTCAGCTGCCAAGACAGCTGT	960
1309	AACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTCTCTACCATTCACCTACAA	1368
961	AACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTCTCTACCATTCACCTACAA	1020
1369	TGGCAGACGCTTCTACTCTCCACCAAGAGGGCGCAGGACGGACATCTTTGGTGCAG	1428
1021	TGGCAGACGCTTCTACTCTCCACCAAGAGGGCGCAGGACGGACATCTTTGGTGCAG	1080
1429	CACAACTTCGAATTTATGAGACAGACCAAGAAATACTCTTTCTGCACAGACCACTGT	1488
1081	CACAACTTCGAATTTATGAGACAGACCAAGAAATACTCTTTCTGCACAGACCACTGT	1140
1489	GGTTTCAGACTCGAGGAGAAATCCAAATGGTGCCTGTGCGCACTTCCCTTCCTATACAA	1548
1141	GGTTTCAGACTCGAGGAGAAATCCAAATGGTGCCTGTGCGCACTTCCCTTCCTATACAA	1200
1549	CAACCAAAATTAACCTGATTGCACTTCTGAGGGCAGAAAGAGACAAATGAGTGGTGTGG	1608
1201	CAACCAAAATTAACCTGATTGCACTTCTGAGGGCAGAAAGAGACAAATGAGTGGTGTGG	1260
1609	GACCACACAGAACTATGATGCGGACCAAGAAAGTTGGGTCTCTGCCCCATGCTGCCACCGA	1668
1261	GACCACACAGAACTATGATGCGGACCAAGAAAGTTGGGTCTCTGCCCCATGCTGCCACCGA	1320
1669	GGAAATCTGCACAAACCAATGAAGGGGTCTGTATCCGCAATGGAGATCAGTGTGATAGCA	1728

1321	db	GGAAATCTGCACCAACCAATGAAGGGTCACTGATCCGATTCGAGATCAGTGGGATAAGCA	1380
1729	QY	GCATGACATGGGTACACATGATGAGGTGCACGTGTGTTGGGAAATGGTCTGGGGAATGGAC	1788
1381	Db	GCATGACATGGGTACACATGATGAGGTGCACGTGTGTTGGGAAATGGTCTGGGGAATGGAC	1440
1789	QY	ATGCATTGCCCTACTCGCAGCTTTCGAGATCAGTGCATTGTTGATGACATCACTTACAAATGT	1848
1441	Db	ATGCATTGCCCTACTCGCACAATTTCGAGATCAGTGCATTGTTGATGACATCACTTACAAATGT	1500
1849	QY	GAAAGACACATTCACAAAGCGCTCATGAAAGAGGGGCACATGCTGAACTGTATCATGCTTCGG	1908
1501	Db	GAAAGACACATTCACAAAGCGCTCATGAAAGAGGGGCACATGCTGAACTGTATCATGCTTCGG	1560
1909	QY	TCAGGGTCGGGGCAGGTGGAAGTGTGATTCGCCGTGCACCAATGCCAGGATTCAGAGACTGG	1968
1561	Db	TCAGGGTCGGGGCAGGTGGAAGTGTGATTCGCCGTGCACCAATGCCAGGATTCAGAGACTGG	1620
1969	QY	GACGTTTATCAAAATTGGAGATTTCATGGGAGAAAGTATGTGTGATGGTGTTCAGATACAGTG	2028
1621	Db	GACGTTTATCAAAATTGGAGATTTCATGGGAGAAAGTATGTGTGATGGTGTTCAGATACAGTG	1680
2029	QY	CTACTGCTATGCGCGTGCCATTGGGGAGTGCATTGGCCAACTTTACAGACCTATCCCAAG	2088
1681	Db	CTACTGCTATGCGCGTGCCATTGGGGAGTGCATTGGCCAACTTTACAGACCTATCCCAAG	1740
2089	QY	CTCAAGTGGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCCGAGCCCACTCCACCC	2148
1741	Db	CTCAAGTGGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCCGAGCCCACTCCACCC	1800
2149	QY	CATCCAGTGGAAATGCACCACAGCCATCTCAATTTCCAAAGTACATTTCTCAGGTGGAGACC	2208
1801	Db	CATCCAGTGGAAATGCACCACAGCCATCTCAATTTCCAAAGTACATTTCTCAGGTGGAGACC	1860
2209	QY	TAAAAATCTGTAGGCCGTTGGAAAGGAGCTTACATACAGGCCACTTTAACTCTCTACAC	2268
1861	Db	TAAAAATCTGTAGGCCGTTGGAAAGGAGCTTACATACAGGCCACTTTAACTCTCTACAC	1920
2269	QY	CATCAAAAGCCCTGAAGCCTGGTGTGGTATACAGAGGCCAGCTCATCAGCATCCAGCAGTA	2328
1921	Db	CATCAAAAGCCCTGAAGCCTGGTGTGGTATACAGAGGCCAGCTCATCAGCATCCAGCAGTA	1980
2329	QY	CGGCCACAAAGAGTGACTCGCTTTGACTTACACACACAGCACCGACACAGCTGTGAC	2388
1981	Db	CGGCCACCAAGAGTGACTCGCTTTGACTTACACACACAGCACCGACACAGCTGTGAC	2040
2389	QY	CAGCAACACCGTGACAGGAGACACACTCCCTTTTCTCTCTTGTGGCCACTTTCTGAATC	2448
2041	Db	CAGCAACACCGTGACAGGAGACACACTCCCTTTTCTCTCTTGTGGCCACTTTCTGAATC	2100
2449	QY	TGTGACCGAAATTCACAGCCAGTAGCTTTGTGTGCTCTGGGTCTCAGCTTCGACACCCGT	2508
2101	Db	TGTGACCGAAATTCACAGCCAGTAGCTTTGTGTGCTCTGGGTCTCAGCTTCGACACCCGT	2160
2509	QY	GTCCGGATTCCGGGTGGAAATATGAGCTGAGTCCGAGGAGGAGATGAGCCACAGTACCCTGGA	2568
2161	Db	GTCCGGATTCCGGGTGGAAATATGAGCTGAGTCCGAGGAGGAGATGAGCCACAGTACCCTGGA	2220
2569	QY	TCTTCCAGGACACGCCACTTCTGTGTGAATATCCCTGACTGCTTCTGGCGCGGAAAATACAT	2628
2221	Db	TCTTCCAGGACACGCCACTTCTGTGTGAATATCCCTGACTGCTTCTGGCGCGGAAAATACAT	2280
2629	QY	TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTGTGATCCTGTCTACTTCACA	2688
2281	Db	TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTGTGATCCTGTCTACTTCACA	2340
2689	QY	AACAAAGCGCCTGATGCGCCCTCTGACCCGACTGTGGACCAAGTGTGATGACACCTCAAT	2748
2341	Db	AACAAAGCGCCTGATGCGCCCTCTGACCCGACTGTGGACCAAGTGTGATGACACCTCAAT	2400
2749	QY	TGTTGTTCCGCTGGAGCAGACCCCAAGGCTCCCATTCACAGGGTACAGAAATAGTCTATTGCC	2808





6781 CTGCACATGCTTTGGAGCCAGCGGGGCTGGCGCTGTGACAACTGCCGCGAGACTGGGGG 6840  
 7189 TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTACCAACCACTATTTCTCAGAGATACCA 7248  
 6841 TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTACCAACCACTATTTCTCAGAGATACCA 6900  
 7249 TCAGAGAACAAACACTAATGTTAATTGTCCTCCCAATGAGTGTCTTCATGCTTTAGATGTACA 7308  
 6901 TCAGAGAACAAACACTAATGTTAATTGTCCTCCCAATGAGTGTCTTCATGCTTTAGATGTACA 6960  
 7309 GGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTTCCCAATCAGAGGAAACAAGCATGT 7368  
 6961 GGCTGACAGAGAGATTCGCCAGAGTAAATCATCTTTCCCAATCAGAGGAAACAAGCATGT 7020  
 7369 CTCTCTCCCAAGATCCATCTAAACTGAGTGTGTAGTGTAGTGTAGTGTAGTGTCTTC 7428  
 7021 CTCTCTCCCAAGATCCATCTAAACTGAGTGTGTAGTGTAGTGTAGTGTCTTC 7080  
 7429 TTCTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGTCAACCTCACAGCT 7488  
 7081 TTCTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGTCAACCTCACAGCT 7140  
 7489 TCTCCAAAGCATACCTCTGGAGTTTCTCTGAGGTTTCTCTAATAAGAGGCTGCACATT 7548  
 7141 TCTCCAAAGCATACCTCTGGAGTTTCTCTGAGGTTTCTCTAATAAGAGGCTGCACATT 7200  
 7549 GCCTGTTCTGCTTCCGAGTATTCATACCGCTCAGTATTTTAAATGAGTGTCTTAAGA 7608  
 7201 GCCTGTTCTGCTTCCGAGTATTCATACCGCTCAGTATTTTAAATGAGTGTCTTAAGA 7260  
 7609 TTGCTTTGGGATCAATAGGAAGCATATGACGACCAACCAAGATGCAAAATGTTTTGAAT 7668  
 7261 TTGCTTTGGGATCAATAGGAAGCATATGACGACCAACCAAGATGCAAAATGTTTTGAAT 7320  
 7669 GATATGACCAAAATTTAAGTAGAAGTACCAACCACTCTGCTTTTCACTAAGTGT 7728  
 7321 GATATGACCAAAATTTAAGTAGAAGTACCAACCACTCTGCTTTTCACTAAGTGT 7380  
 7729 CTGGCCCGCAATCTGTAGGAACCAAGCATGATCTTGTGTGATATTTTAAATATCCA 7788  
 7381 CTGGCCCGCAATCTGTAGGAACCAAGCATGATCTTGTGTGATATTTTAAATATCCA 7440  
 7789 CAGTACT 7795  
 7441 CAGTACT 7447

RESULT 11  
 ACF03878  
 ID ACF03878 standard; cDNA; 7680 BP.  
 XX  
 AC ACF03878;  
 XX  
 DT 18-SEP-2003 (first entry)  
 XX  
 DE Human fibronectin encoding cDNA SEQ ID NO:9.

Human; integrin alpha4beta1; fibronectin; integrin alpha 4; VCAM;  
 integrin beta 1; vascular cell adhesion molecule; angiogenesis;  
 integrin alpha4beta1 binding inhibitor; angiogenesis inhibitor;  
 cyostatic; ophthalmological; antiarthritic; cancer;  
 antiinflammatory; osteopathic; dermatological; gene therapy; psoriasis;  
 pathological condition; diabetic retinopathy; macular degeneration;  
 neovascularisation; rheumatoid arthritis; osteoarthritis; skin cancer;  
 endothelial progenitor cell; wound healing; gene; ss.

Homo sapiens.

WO2003019136-A2.

06-MAR-2003.

01-AUG-2002; 2002WO-US024573.

XX 06-AUG-2001; 2001US-0310645P.  
 PR (REGC ) UNIV CALIFORNIA.  
 XX Varner JA;  
 XX WPI; 2003-278690/27.  
 DR P-PSDB; ABR81866.  
 XX  
 PT Inhibiting angiogenesis in a tissue, useful for treating cancer,  
 PT arthritis, retinopathy, psoriasis, by providing a tissue and an agent  
 PT that inhibits specific binding of integrin alpha-4beta-1 to an integrin  
 PT alpha-4beta-1 ligand.  
 XX  
 PS Disclosure; Fig 17; 177pp; English.  
 XX  
 CC The present invention describes a method for inhibiting angiogenesis in a  
 CC tissue comprising providing a tissue and an agent that inhibits specific  
 CC binding of integrin alpha4beta1 to an integrin alpha4beta1 ligand, and  
 CC treating the tissue with the agent under conditions where the specific  
 CC binding of the alpha4beta1 to the ligand is inhibited and a treated  
 CC tissue is produced. Angiogenesis in the treated tissue is inhibited. Also  
 CC described: (1) inhibiting endothelial cell adhesion or migration; (2)  
 CC detecting angiogenesis in a tissue; (3) screening a test compound; (4)  
 CC isolating endothelial progenitor cells from a tissue; and (5) reducing  
 CC symptoms associated with cancer in a subject, or a pathological condition  
 CC is an ocular or skin tissue. An integrin alpha4beta1 binding inhibitor  
 CC has cytostatic, ophthalmological, antiarthritic, antiinflammatory,  
 CC antineoplastic, osteopathic and dermatological activities, and can be  
 CC used in gene therapy. The methods are useful for treating cancer, and  
 CC other pathological condition, such as diabetic retinopathy, macular  
 CC degeneration by neovascularisation, rheumatoid arthritis, osteoarthritis,  
 CC psoriasis or skin cancer. The methods are also useful in isolating  
 CC endothelial progenitor cells, and in determining the mechanisms that  
 CC underlie angiogenesis, development, wound healing and the function of the  
 CC female reproductive system. The present sequence is used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Query Match 95.2%; Score 7423; DB 7; Length 7680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 349 GAAGAGCAAGAGGCGAGGCTCAGCAAAATGGTTACGCCCCAGTCCCGGTGGTGTCA 408  
 DB 1 GAAGAGCAAGAGGCGAGGCTCAGCAAAATGGTTACGCCCCAGTCCCGGTGGTGTCA 60  
 QY 409 AAGCAAGCCCGTGTGTATGACATGGAACACCTATCAGATAAATCAACAGTGGAGCG 468  
 DB 61 AAGCAAGCCCGTGTGTATGACATGGAACACCTATCAGATAAATCAACAGTGGAGCG 120  
 QY 469 GACCTACCTAGGCAATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 528  
 DB 121 GACCTACCTAGGCAATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
 QY 529 CGAAGTAAACCTGAAGTGAAGAGACTTGTTCACAACTACACTGGACACTTACCG 588  
 DB 181 CGAAGTAAACCTGAAGTGAAGAGACTTGTTCACAACTACACTGGACACTTACCG 240  
 QY 589 AGTGGGTGACACTTATGAGCGTCTTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGG 648  
 DB 241 AGTGGGTGACACTTATGAGCGTCTTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGG 300  
 QY 649 GGCTGGGCGAGGAGAAATAGCTGTACCATGTGCAAAACCGCTGCCATGAAGGGGGTCAGTC 708  
 DB 301 GGCTGGGCGAGGAGAAATAGCTGTACCATGTGCAAAACCGCTGCCATGAAGGGGGTCAGTC 360  
 QY 709 CTACAAGATTGGTGTGACACCTGGAGGAGAGCCACATGAGACTGGTGTGTGTGTGTGTGT 768  
 DB 361 CTACAAGATTGGTGTGACACCTGGAGGAGAGCCACATGAGACTGGTGTGTGTGTGTGTGT 420







QY 5149 TGCAGTAACCAACATGATCGCCCTAAAGGAGTGGCAATTCATCTGATGTGGATGTCGATTC 5208  
 Db 4801 TGCAGTAACCAACATGATCGCCCTAAAGGAGTGGCAATTCATCTGATGTGGATGTCGATTC 4860  
 QY 5209 CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTACCTACTC 5268  
 Db 4861 CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTACCTACTC 4920  
 QY 5269 GAGCCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACCTGATGTGTAAGAGACACTGC 5328  
 Db 4921 GAGCCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACCTGATGTGTAAGAGACACTGC 4980  
 QY 5329 AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGAGTGTGCTGTCAGCA 5388  
 Db 4981 AGAGCTGCAAGGCTCAGACCGGTTCTGAGTACACAGTCACTGAGTGTGCTGTCAGCA 5040  
 QY 5389 TGAATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5448  
 Db 5041 TGAATGAGAGCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCCTGACCAACTGA 5100  
 QY 5449 CTTGAAGTTCACTCAGGTCAACCCCAAGCCTGAGCGGCCAGTGGACACCAACCCTGAT 5508  
 Db 5101 CTTGAAGTTCACTCAGGTCAACCCCAAGCCTGAGCGGCCAGTGGACACCAACCCTGAT 5160  
 QY 5509 TCAGCTCACTGAGATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568  
 Db 5161 TCAGCTCACTGAGATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220  
 QY 5569 AATCAACCTTGTCTGAGCAGCTCACTGCTGTTGTTGATCAGGACTTATGTTGGCCACCAA 5628  
 Db 5221 AATCAACCTTGTCTGAGCAGCTCACTGCTGTTGTTGATCAGGACTTATGTTGGCCACCAA 5280  
 QY 5629 ATATGAAGTGAAGTCTATGCTCTTAAGGACACTTTGACACGACAGCAGCTCAGGGAGT 5688  
 Db 5281 ATATGAAGTGAAGTCTATGCTCTTAAGGACACTTTGACACGACAGCAGCTCAGGGAGT 5340  
 QY 5689 TGTCAACCTTGTGAGAAATGTCAAGCCCAAGAGGCTGCTGTGACAGATGCTACTGA 5748  
 Db 5341 TGTCAACCTTGTGAGAAATGTCAAGCCCAAGAGGCTGCTGTGACAGATGCTACTGA 5400  
 QY 5749 GACCAATCACTAGCTTGGAGACCAAGTCAAGTCACTGAGCAGTCACTGGCCCTCCAGTTGA 5808  
 Db 5401 GACCAATCACTAGCTTGGAGACCAAGTCAAGTCACTGAGCAGTCACTGGCCCTCCAGTTGA 5460  
 QY 5809 TGCAGTTCAGGCAATGAGGCTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5868  
 Db 5461 TGCAGTTCAGGCAATGAGGCTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAG 5520  
 QY 5869 CTAACCAATCACTGGCTTCAACAGGCACTGACTACAGATCTACCTGTACACCTTGA 5928  
 Db 5521 CTAACCAATCACTGGCTTCAACAGGCACTGACTACAGATCTACCTGTACACCTTGA 5580  
 QY 5929 TGAATGTCTGGAGCTCCCTGTGCTCATCGAGCCTCCACTGCTGATTTGATGCACTATC 5988  
 Db 5581 TGAATGTCTGGAGCTCCCTGTGCTCATCGAGCCTCCACTGCTGATTTGATGCACTATC 5640  
 QY 5989 CAACCTGCGTTTCTGGCCACCAACCCAAATTCCTTGTGCTGATCATGAGCAGCCGACG 6048  
 Db 5641 CAACCTGCGTTTCTGGCCACCAACCCAAATTCCTTGTGCTGATCATGAGCAGCCGACG 5700  
 QY 6049 TGCAGGATTAACCGCTACATCATCAAGTATGAGAGCCTGCTGCTCCAGAGAGT 6108  
 Db 5701 TGCAGGATTAACCGCTACATCATCAAGTATGAGAGCCTGCTGCTCCAGAGAGT 5760  
 QY 6109 GGTCCCTCGGCCCGCTGCTGCTCAAGAGTCTATTAATGAGCCTGGAACCCGGGAC 6168  
 Db 5761 GGTCCCTCGGCCCGCTGCTGCTCAAGAGTCTATTAATGAGCCTGGAACCCGGGAC 5820  
 QY 6169 CGAATATCAATTTATGCTATTCCTGAGATATCAAGAGCAGGCGCCCTGATTTGG 6228  
 Db 5821 CGAATATCAATTTATGCTATTCCTGAGATATCAAGAGCAGGCGCCCTGATTTGG 5880  
 QY 6229 AAGGAAAAAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCACTGG 6288

Db 5881 AAGGAAAAAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCACTGG 5940  
 QY 6289 ACCAGAGATCTTGGATGTTTCTTCCACAGTCAAAAGACCCCTTTCGTACCCACCCCTGG 6348  
 Db 5941 ACCAGAGATCTTGGATGTTTCTTCCACAGTCAAAAGACCCCTTTCGTACCCACCCCTGG 6000  
 QY 6349 GTATGACATCTGGAATGATTTTCAGCTTCTGGCACTTCTGTGTCAGCAACCCAGTGTGG 6408  
 Db 6001 GTATGACATCTGGAATGATTTTCAGCTTCTGGCACTTCTGTGTCAGCAACCCAGTGTGG 6060  
 QY 6409 GCAACAAATGATCTTTGAGGAACATGCTTTTAGGCGGACCAACCCGCCCAACCCGCCAC 6468  
 Db 6061 GCAACAAATGATCTTTGAGGAACATGCTTTTAGGCGGACCAACCCGCCCAACCCGCCAC 6120  
 QY 6469 CCCATAAAGCATAGGCCCAAGACCATACCCCGGGAATAGAGCAAGAGCTCTCTCA 6528  
 Db 6121 CCCATAAAGCATAGGCCCAAGACCATACCCCGGGAATAGAGCAAGAGCTCTCTCA 6180  
 QY 6529 GACAACTCTCATGAGGCGCCCATTTCCAGGACACTTCTGAGTACATCATTTTCACTGTCTCC 6588  
 Db 6181 GACAACTCTCATGAGGCGCCCATTTCCAGGACACTTCTGAGTACATCATTTTCACTGTCTCC 6240  
 QY 6589 TGTGGCACTGATGAAGAACCCCTTACAGTTCAGGTTCTGGAACCTTCTTACAGTGCAC 6648  
 Db 6241 TGTGGCACTGATGAAGAACCCCTTACAGTTCAGGTTCTGGAACCTTCTTACAGTGCAC 6300  
 QY 6649 TCTGACAGGCTCACCAGAGGTGCTCCACTACACATCATAGTGGAGGCACTGGAAGCCA 6708  
 Db 6301 TCTGACAGGCTCACCAGAGGTGCTCCACTACACATCATAGTGGAGGCACTGGAAGCCA 6360  
 QY 6709 GCAGAGCATTAAGGTTTCGGGAAGAGGTTGTTACCTGGGCAACTCTGTCAAGAGGCTT 6768  
 Db 6361 GCAGAGCATTAAGGTTTCGGGAAGAGGTTGTTACCTGGGCAACTCTGTCAAGAGGCTT 6420  
 QY 6769 GAACCAACCTACCGATGACTGCTTGGACCTTACAGTGTCCCATTTATCCCGTTGG 6828  
 Db 6421 GAACCAACCTACCGATGACTGCTTGGACCTTACAGTGTCCCATTTATCCCGTTGG 6480  
 QY 6829 AGATGAGTGGGAACGATGTCTGAATCAGGCTTTAAACTGTTGTCAGTGTGCTT 6888  
 Db 6481 AGATGAGTGGGAACGATGTCTGAATCAGGCTTTAAACTGTTGTCAGTGTGCTT 6540  
 QY 6889 TGGAGTGTCTATTCAGATGTGATTTCACTAGATGCTGTCATGACATGCTGTAAC 6948  
 Db 6541 TGGAGTGTCTATTCAGATGTGATTTCACTAGATGCTGTCATGACATGCTGTAAC 6600  
 QY 6949 CAAGATTGGAGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATG 7008  
 Db 6601 CAAGATTGGAGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATG 6660  
 QY 7009 TCTTGGGAAACGGGAAAGAGAAATTCAGTGTGACCTCATGAGGCAACGTTTACGATGA 7068  
 Db 6661 TCTTGGGAAACGGGAAAGAGAAATTCAGTGTGACCTCATGAGGCAACGTTTACGATGA 6720  
 QY 7069 TGGGAACACATACCACTAGGAGAACAGTGGCAGAGGAAATATCTCGGTGCCATTTGCTC 7128  
 Db 6721 TGGGAACACATACCACTAGGAGAACAGTGGCAGAGGAAATATCTCGGTGCCATTTGCTC 6780  
 QY 7129 CTGCACATGCTTTGGAGGCGCAGGGGCTGCGCTGTGACAACTGCGGCAAGACCTGGGG 7188  
 Db 6781 CTGCACATGCTTTGGAGGCGCAGGGGCTGCGCTGTGACAACTGCGGCAAGACCTGGGG 6840  
 QY 7189 TGAACCCAGTCCCGAGGCACTACTGCGCAGTCTCAACCCAGTATTCTCAGAGATACCA 7248  
 Db 6841 TGAACCCAGTCCCGAGGCACTACTGCGCAGTCTCAACCCAGTATTCTCAGAGATACCA 6900  
 QY 7249 TCAGAGAACAAACACTAATGTTAATTTGCCCAATGAGTGTCTTCACTGCTTTAGATGPACA 7308  
 Db 6901 TCAGAGAACAAACACTAATGTTAATTTGCCCAATGAGTGTCTTCACTGCTTTAGATGPACA 6960  
 QY 7309 GGTCTGACAGAGAGATTCGCGAGAGTAAATCATCTTCCCAATCCAGAGGAAACAGCATGT 7368

2b 6961 GGCTGACAGAGAGATTCGCGAGTAGTAATCATCTTTCCATCCAGAGGAACAGCATGT 7020  
 2y CTCTCTCCAGAGATTCATTAACCTGAGTAGTAGTAGCAGACCCAGCTTAGAGTTCTTC 7428  
 2b CTCTCTCCAGAGATTCATTAACCTGAGTAGTAGTAGCAGACCCAGCTTAGAGTTCTTC 7080  
 2y TTCTCTTCTTAAGCCCTTCTCTGAGGAGTAGTCTCCAGCTTCAGCTCAACTCAGCT 7488  
 2b TTCTCTTCTTAAGCCCTTCTCTGAGGAGTAGTCTCCAGCTTCAGCTCAACTCAGCT 7140  
 2y TCTCCAGCATCACCTCTGGAGTTCTCTGAGGTTCTCTATAAATGAGGCTTGCACTT 7548  
 2b TCTCCAGCATCACCTCTGGAGTTCTCTGAGGTTCTCTATAAATGAGGCTTGCACTT 7200  
 2y GCCTGTTCTCTGAGTAGTATCAATACCCCTCAGTATTTTAAATGAGGTATCTTAA 7608  
 2b GCCTGTTCTCTGAGTAGTATCAATACCCCTCAGTATTTTAAATGAGGTATCTTAA 7260  
 2y TTTGGTTTGGATCAATPAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTCAAT 7668  
 2b TTTGGTTTGGATCAATPAGGAAGCATATGAGCAACCAAGATGCAAAATGTTTCAAT 7320  
 2y GATATGACCAAAATTTAAGTAGGAAGTCAACCAACCACTTCTGCTTTCACCTAAAGT 7728  
 2b GATATGACCAAAATTTAAGTAGGAAGTCAACCAACCACTTCTGCTTTCACCTAAAGT 7380  
 2y CTGSCCCGCAATCTGTAGGAAGCATATGCTTGTCTGTATTTTAAATATATCA 7788  
 2b CTGSCCCGCAATCTGTAGGAAGCATATGCTTGTCTGTATTTTAAATATATCA 7440  
 2y CAGTACT 7795  
 2b CAGTACT 7447

RESULT 12

ABX10391  
 ID ABX10391 standard; DNA; 7680 BP.  
 AC AC  
 CX ABX10391;  
 DT 28-JAN-2003 (first entry)  
 DX DNA encoding protein differentially regulated in prostate cancer #60.  
 DE Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;  
 KW Gene; ds.  
 KX Homo sapiens.  
 DS WO200281638-A2.  
 FN 17-OCT-2002.  
 PD 08-APR-2002; 2002WO-US010824.  
 PF 06-APR-2001; 2001US-0281731P.  
 PR 06-APR-2001; 2001US-0281732P.  
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
 PA Sun Z, Jay G;  
 PI WPI; 2003-058520/05.  
 DR P-PSDB; ABU07486.  
 XX Novel genes which are differentially regulated in prostate cancer, useful  
 PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 PT therapeutic or preventive intervention in prostate cancer patients.  
 XX Claim 1; Page 136-139; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in  
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 CC sample comprising prostate tissue, which involves determining the number  
 CC of target genes which are differentially-regulated in the sample, where  
 CC the number is indicative of the probability that the sample comprises  
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 CC intervention in a subject having a prostate cancer, which involves  
 CC determining the expression levels in a sample comprising prostate tissue  
 CC of target genes which are differentially-regulated in prostate cancer.  
 CC Preferably, the expression levels of at least 10 genes are determined.  
 CC (I) is also useful for identifying agents that modulate a biological  
 CC activity of a polypeptide differentially-regulated in prostate cancer  
 CC cells, which involves contacting a polypeptide differentially-regulated  
 CC in prostate cancer cells with a test agent under conditions effective for  
 CC the test agent to modulate a biological activity of the polypeptide, and  
 CC determining whether the test agent modulates the biological activity. (I)  
 CC is useful as molecular markers, as drug targets, and for detecting,  
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 CC preventing or treating, determining predisposition to diseases and  
 CC conditions especially relating to prostate cancer. (I) and its expression  
 CC products are used in the diagnostic test to assay for presence of cancer  
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 CC of cancer, its stage of development, the nature of genetic defect, etc.  
 CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 CC discovery. (I) can also be used for expressing the polypeptide and thus  
 CC for searching specific binding partners of the polypeptide. (I) is useful  
 CC in therapeutic applications to treat prostate cancer. The identification  
 CC of specific genes, and groups of genes, expressed in pathways  
 CC physiologically relevant to prostate cancer permits the definition of  
 CC functional and disease pathways and the delineation of targets in these  
 CC pathways which are useful in diagnostic, therapeutic, and clinical  
 CC applications. This sequence encodes a protein differentially regulated in  
 CC prostate cancer  
 XX  
 SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;  
 Query Match 95.2%; Score 7423; DB 7; Length 7680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 349 GAAGAGCAGAGGAGGCTCAGCAATGTTTTCAGCCAGTCCCGTGGTGTTCAGTCA 408  
 DB 1 GAAGAGCAGAGGAGGCTCAGCAATGTTTTCAGCCAGTCCCGTGGTGTTCAGTCA 60  
 QY 409 AAGCAAGCCCGTGTCTTATGCAATGGAACAACTATCAGATAAATCAACAGTGGAGCG 468  
 DB 61 AAGCAAGCCCGTGTCTTATGCAATGGAACAACTATCAGATAAATCAACAGTGGAGCG 120  
 QY 469 GACCTACCTAGGCATGCTTGTGTTTGTACTTGTATGAGGAGCCCGAGTTTAACTG 528  
 DB 121 GACCTACCTAGGCATGCTTGTGTTTGTACTTGTATGAGGAGCCCGAGTTTAACTG 180  
 QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCG 588  
 DB 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCG 240  
 QY 589 AGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTCTACTGTCATCGG 648  
 DB 241 AGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTCTACTGTCATCGG 300  
 QY 649 GGCTGGGCGAGGAGGAGTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCACTC 708  
 DB 301 GGCTGGGCGAGGAGGAGTAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGGTCACTC 360  
 QY 709 CTACAAGATTGGTGACACTGAGGAGGAGCCATGAGACTGGTGTGTACATGTTAGAGTG 768  
 DB 361 CTACAAGATTGGTGACACTGAGGAGGAGCCATGAGACTGGTGTGTGTACATGTTAGAGTG 420  
 QY 769 TGTGTGCTCTTGTGTATGGAAGAGAGATGACCTGCAAGCCCATAGCTGAGAAGTGT 828  
 DB 421 TGTGTGCTCTTGTGTATGGAAGAGAGATGACCTGCAAGCCCATAGCTGAGAAGTGT 480





Db 2641 |||||GCCCTCTCCAGGACCTGCAAGTTTGTGGAGTGCACAGAGCTCAACATCATGTG 2700  
Qy 3049 GACACCGCTCAGAGTGCAGTGACCGGCTACCGTGTGGATGTGATCCCGCTCAACCTGCC 3108  
Db 2701 GACACCGCTCAGAGTGCAGTGACCGGCTACCGTGTGGATGTGATCCCGCTCAACCTGCC 2760  
Qy 3109 TGGCGAGCAGCGGAGAGGCTGCCATCAGCAGAAACACCTTTGCGAAGTGCACCGGCT 3168  
Db 2761 TGGCGAGCAGCGGAGAGGCTGCCATCAGCAGAAACACCTTTGCGAAGTGCACCGGCT 2820  
Qy 3169 GTCCCTGGGTCACTTACTTCAAGTCTTTGAGTGCAGCCATCGGAGGAGAGCAA 3228  
Db 2821 GTCCCTGGGTCACTTACTTCAAGTCTTTGAGTGCAGCCATCGGAGGAGAGCAA 2880  
Qy 3229 CCCTCTGACTCTCAACAGACAAACAACTGGATGCTCCCACTAACTCAAGTTGTCAA 3288  
Db 2881 CCCTCTGACTCTCAACAGACAAACAACTGGATGCTCCCACTAACTCAAGTTGTCAA 2940  
Qy 3289 TGAACCTGATTTACTGCTGCTGAGTGGACTCCACCTCGGCGCCAGATACAGGATA 3348  
Db 2941 TGAACCTGATTTACTGCTGCTGAGTGGACTCCACCTCGGCGCCAGATACAGGATA 3000  
Qy 3349 CCGACTGACCGTGGGCTTTACCCGAAAGAGCAGCCAGGAGTACAAATGTGGTCCCTC 3408  
Db 3001 CCGACTGACCGTGGGCTTTACCCGAAAGAGCAGCCAGGAGTACAAATGTGGTCCCTC 3060  
Qy 3409 TGTCTCAAGTACCACTGAGGATCTGCAGCTGCTGATCTGAGTACACCGTATCCCTCGT 3468  
Db 3061 TGTCTCAAGTACCACTGAGGATCTGCAGCTGCTGATCTGAGTACACCGTATCCCTCGT 3120  
Qy 3469 GGCATAAAGGCGCAACCAAGAGAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3528  
Db 3121 GGCATAAAGGCGCAACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180  
Qy 3529 TGGAGTCTTATTCACCTTACACCGAGGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGG 3588  
Db 3181 TGGAGTCTTATTCACCTTACACCGAGGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGG 3240  
Qy 3589 GACGCTCTCCAGAAATGTTTAACTGGGTGACGACCAAGGAGGAGGAGGAGGAGGAGGAGG 3648  
Db 3241 GACGCTCTCCAGAAATGTTTAACTGGGTGACGACCAAGGAGGAGGAGGAGGAGGAGGAGG 3300  
Qy 3649 ACCAGAGAGTGAATTCAGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGT 3708  
Db 3301 ACCAGAGAGTGAATTCAGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGT 3360  
Qy 3709 AGAATACCTTACACCTTCCAGTCCAGAGATGGAAGAGAGATGCGCCAAATGT 3768  
Db 3361 AGAATACCTTACACCTTCCAGTCCAGAGATGGAAGAGAGATGCGCCAAATGT 3420  
Qy 3769 AAACAAAGTGTGACACCAATGTCTCCACCAACAACTTGATCTGAGGAGCAACCTTGA 3828  
Db 3421 AAACAAAGTGTGACACCAATGTCTCCACCAACAACTTGATCTGAGGAGCAACCTTGA 3480  
Qy 3829 CACTGGAGTGTCAAGTCTCTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3888  
Db 3481 CACTGGAGTGTCAAGTCTCTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540  
Qy 3889 AATTACCAACACCCCTACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3948  
Db 3541 AATTACCAACACCCCTACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600  
Qy 3949 TGATCAGAGTCTCTGACCTTTTATTAACCTGAGTCCCGGCTCGAGTACAAATGTCACTGT 4008  
Db 3601 TGATCAGAGTCTCTGACCTTTTATTAACCTGAGTCCCGGCTCGAGTACAAATGTCACTGT 3660  
Qy 4009 TTACACTGTCAAGATGACAAAGAAAGTGTCCCTATCTCTGATACCACTCATCCAGCTGT 4068  
Db 3661 TTACACTGTCAAGATGACAAAGAAAGTGTCCCTATCTCTGATACCACTCATCCAGCTGT 3720  
Qy 4069 TCCTCTCCACTGACCTGCGAATCCCAACATTTGGTCCAGACCCAGTGCCTGTCACTGT 4128

Db 3721 TCCTCTCCACTGACCTGCGAATCCAAACATTTGGTCCAGACCATGCGGTGTCACTGT 3780  
Qy 4129 GGCTCCACCCCATCATTTGATTTAACCAACTTCTCGTGGTGTACTCACCTGTGAAAAA 4188  
Db 3781 GGCTCCACCCCATCATTTGATTTAACCAACTTCTCGTGGTGTACTCACCTGTGAAAAA 3840  
Qy 4189 TGAGGAAGTGTGAGAGTGTGCAATTTCTCTCTCAGACATGAGTGTCTTAAACAA 4248  
Db 3841 TGAGGAAGTGTGAGAGTGTGCAATTTCTCTCTCAGACATGAGTGTCTTAAACAA 3900  
Qy 4249 TCTCTGCTGCTGATACAGAAATATGATGAGTGTCTCCAGTGTCTACGAAACATGAGAG 4308  
Db 3901 TCTCTGCTGCTGATACAGAAATATGATGAGTGTCTCCAGTGTCTACGAAACATGAGAG 3960  
Qy 4309 CACACTCTTAGAGAGACAGAAACAGGCTTGAATTTCCCAACTGGCATTTTC 4368  
Db 3961 CACACTCTTAGAGAGACAGAAACAGGCTTGAATTTCCCAACTGGCATTTTC 4020  
Qy 4369 TGATATTACTGCAACTCTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCACTGG 4428  
Db 4021 TGATATTACTGCAACTCTTTTACTGTGCACTGGATTTGCTCTCGAGCCACCATCACTGG 4080  
Qy 4429 CTACAGGATCCGCAATCATCCGAGCACTTCACTGGGAGACCTCGAGAAAGATCGGTTGCC 4488  
Db 4081 CTACAGGATCCGCAATCATCCGAGCACTTCACTGGGAGACCTCGAGAAAGATCGGTTGCC 4140  
Qy 4489 CCACTCTCGGAATTCATCACCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAG 4548  
Db 4141 CCACTCTCGGAATTCATCACCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAG 4200  
Qy 4549 CATGCTGTCTTAAATGGCAGAGAGAAAGTCCCTTATTGATTTGGCCCAACAACTCAAGT 4608  
Db 4201 CATGCTGTCTTAAATGGCAGAGAGAAAGTCCCTTATTGATTTGGCCCAACAACTCAAGT 4260  
Qy 4609 TTCTGATTTCCGAGGAGCTCGGAAGTTGTTCTCGGACCCCAACAGCTCTGATCAG 4668  
Db 4261 TTCTGATTTCCGAGGAGCTCGGAAGTTGTTCTCGGACCCCAACAGCTCTGATCAG 4320  
Qy 4669 CTGGGATGCTCTGCTGTCACTGAGATATTACAGGATCACTTACGAGAAACAGGAGG 4728  
Db 4321 CTGGGATGCTCTGCTGTCACTGAGATATTACAGGATCACTTACGAGAAACAGGAGG 4380  
Qy 4729 AAATAGCCCTGTCCAGGAGTTCACCTGCTGGGAGCAAGTCTACAGTACCATCAGGG 4788  
Db 4381 AAATAGCCCTGTCCAGGAGTTCACCTGCTGGGAGCAAGTCTACAGTACCATCAGGG 4440  
Qy 4789 CCTTAAACCTGGAGTTGATTAACCATCACTGTGTATGTCTCACTGGCCGTGGAGACAG 4848  
Db 4441 CCTTAAACCTGGAGTTGATTAACCATCACTGTGTATGTCTCACTGGCCGTGGAGACAG 4500  
Qy 4849 CCCGCAAGCAGCAAGCCAAATTTCAATTTACCGAACAGAAATGACAAACCATCCCA 4908  
Db 4501 CCCGCAAGCAGCAAGCCAAATTTCAATTTACCGAACAGAAATGACAAACCATCCCA 4560  
Qy 4909 GATGCAAGTGAACCGATTTTCAGGACAAACAGCAATAGTGTCAAGTGTGCTTCAAGTTTC 4968  
Db 4561 GATGCAAGTGAACCGATTTTCAGGACAAACAGCAATAGTGTCAAGTGTGCTTCAAGTTTC 4620  
Qy 4969 CCTGTTACTGTGTTACAGAGTAAACCACTCCCAAAAATGACAGGAGCAACAAACAAAC 5028  
Db 4621 CCTGTTACTGTGTTACAGAGTAAACCACTCCCAAAAATGACAGGAGCAACAAACAAAC 4680  
Qy 5029 TAAACTGCAGTCCAGATCAACAGAAATGACTATTGAAGGCTTGCAGCCACAGTGA 5088  
Db 4681 TAAACTGCAGTCCAGATCAACAGAAATGACTATTGAAGGCTTGCAGCCACAGTGA 4740  
Qy 5089 GTATGTGTTAGTGTCTATGCTCAGAAATCCAAAGCAGAGATGACGCTCTGTTGAGAC 5148  
Db 4741 GTATGTGTTAGTGTCTATGCTCAGAAATCCAAAGCAGAGATGACGCTCTGTTGAGAC 4800  
Qy 5149 TGCAGTAAACCAATTTGATCGGCTTAAAGGATTCGCAATTCAGTGTGATGTGATTC 5208  
Db 4801 TGCAGTAAACCAATTTGATCGGCTTAAAGGATTCGCAATTCAGTGTGATGTGATTC 4860

QY 5209 CATCAAAATTGCTGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGTTGACCTACTC 5268  
 Db 4861 CATCAAAATTGCTGGAAAGCCACAGGGCAAGTTTCCAGGTACAGGTTGACCTACTC 4920  
 QY 5269 GAGCCCTGAGGATGGAATCCATGAGATTAATCCCTGCACCTGATGTTGAAGAAGACACTGC 5328  
 Db 4921 GAGCCCTGAGGATGGAATCCATGAGATTAATCCCTGCACCTGATGTTGAAGAAGACACTGC 4980  
 QY 5329 AGAGCTGCAAGGCCCTCAGACCGGGTTCTGAGTACACAGTCAAGTGTGCTTGCACGA 5388  
 Db 4981 AGAGCTGCAAGGCCCTCAGACCGGGTTCTGAGTACACAGTCAAGTGTGCTTGCACGA 5040  
 QY 5389 TGATATGAGAGGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTTGCAACCACTGA 5448  
 Db 5041 TGATATGAGAGGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTTGCAACCACTGA 5100  
 QY 5449 CTTGAAGTTCACTCAGGTACACCCACACAGCTGAGCGCCAGTGCACACCAACCAATGT 5508  
 Db 5101 CTTGAAGTTCACTCAGGTACACCCACACAGCTGAGCGCCAGTGCACACCAACCAATGT 5160  
 QY 5509 TCAGCTCACTGGATATCGAGTGCAGGTGACCCCAAGGAGAACCCGACCAATGAAGA 5568  
 Db 5161 TCAGCTCACTGGATATCGAGTGCAGGTGACCCCAAGGAGAACCCGACCAATGAAGA 5220  
 QY 5569 AATCAACCTTGCTCTGACAGCTATCCCTGTTGTTATCAGGACTTATGTTGGCCACCAA 5628  
 Db 5221 AATCAACCTTGCTCTGACAGCTATCCCTGTTGTTATCAGGACTTATGTTGGCCACCAA 5280  
 QY 5629 ATATGAAGTGAGTGTCTATGCTCTTAAGGACACTTTGACACGACAGCAGCTCAGGGAGT 5688  
 Db 5281 ATATGAAGTGAGTGTCTATGCTCTTAAGGACACTTTGACACGACAGCAGCTCAGGGAGT 5340  
 QY 5689 TGTCAACACTTGGAGAAATGTGACCCCAACAGAGGGCTCTGTGTACAGATGCTACTGA 5748  
 Db 5341 TGTCAACACTTGGAGAAATGTGACCCCAACAGAGGGCTCTGTGTACAGATGCTACTGA 5400  
 QY 5749 GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCCCTCCAAAGTTGA 5808  
 Db 5401 GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCCCTCCAAAGTTGA 5460  
 QY 5809 TGCCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAACCCATCAAGCCAGATGTCAGAG 5868  
 Db 5461 TGCCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAACCCATCAAGCCAGATGTCAGAG 5520  
 QY 5869 CTACACCATCAGTGGCTTACACCCAGSACTGACTACAGATCTACTCTGACACCTTGAA 5928  
 Db 5521 CTACACCATCAGTGGCTTACACCCAGSACTGACTACAGATCTACTCTGACACCTTGAA 5580  
 QY 5929 TGACAAATGCTCGGAGCTCCCTGTGTTGATCGACGCTCCACTGCGAATGATGCAACATC 5988  
 Db 5581 TGACAAATGCTCGGAGCTCCCTGTGTTGATCGACGCTCCACTGCGAATGATGCAACATC 5640  
 QY 5989 CACCTGCGTTCTCGGCCACACACCCAAATTCCTGCTGTTATCATGCGCCGCCACG 6048  
 Db 5641 CACCTGCGTTCTCGGCCACACACCCAAATTCCTGCTGTTATCATGCGCCGCCACG 5700  
 QY 6049 TGCAGGATTAACCGGCTACATCATCAAGTATGAGAACCTGGGTCTCTCCCCAGAGAAAT 6108  
 Db 5701 TGCAGGATTAACCGGCTACATCATCAAGTATGAGAACCTGGGTCTCTCCCCAGAGAAAT 5760  
 QY 6109 GGTCCCTCGGCCCGCCCTCGGTGTACAGAGGCTACTATTACTGGCCCTGGACCCGGAAC 6168  
 Db 5761 GGTCCCTCGGCCCGCCCTCGGTGTACAGAGGCTACTATTACTGGCCCTGGACCCGGAAC 5820  
 QY 6169 CGAATATACAAATTTATGCTATTGCTCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTGG 6228  
 Db 5821 CGAATATACAAATTTATGCTATTGCTCCCTGAAGAAATATCAGAGAGCGAGCCCTGATTGG 5880  
 QY 6229 AAGGAAAAAGACAGAGCTTCCCAACCTGGTAACTCCCTGACACCCCAATCTTCATGG 6288  
 Db 5881 AAGGAAAAAGACAGAGCTTCCCAACCTGGTAACTCCCTGACACCCCAATCTTCATGG 5940

QY 6289 ACCAGAGATCTTGGATGTTCTTCCATCCAGTTCACAGTTCACAAAGACCCCTTCTGTCACCCACCTGG 6348  
 Db 5941 ACCAGAGATCTTGGATGTTCTTCCATCCAGTTCACAAAGACCCCTTCTGTCACCCACCTGG 6000  
 QY 6349 GTATGACACTGGAAATGTTATTCAGCTTCTGCTGCTCAGTCAAGAACCCAGTGTGG 6408  
 Db 6001 GTATGACACTGGAAATGTTATTCAGCTTCTGCTGCTCAGTCAAGAACCCAGTGTGG 6060  
 QY 6409 GCAACAAATGATCTTTGAGGAACATGTTTATAGCGGACCAACACCGCCCAACAGGCCAC 5468  
 Db 6061 GCAACAAATGATCTTTGAGGAACATGTTTATAGCGGACCAACACCGCCCAACAGGCCAC 6120  
 QY 6469 CCCCATAGGCTAGGCCAAGACCATACCGCGGATGTAGGACAAAGAGCTCTCTCTCA 6528  
 Db 6121 CCCCATAGGCTAGGCCAAGACCATACCGCGGATGTAGGACAAAGAGCTCTCTCTCA 6180  
 QY 6529 GACAACCATCTCATGGGCCCCCAATTCAGGACACTTCTGAGTACATCATTTTCAATCTC 6588  
 Db 6181 GACAACCATCTCATGGGCCCCCAATTCAGGACACTTCTGAGTACATCATTTTCAATCTC 6240  
 QY 6589 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGGTTCTGGAACTTCTTACCAAGTCCAC 6648  
 Db 6241 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGGTTCTGGAACTTCTTACCAAGTCCAC 6300  
 QY 6649 TCTGACAGGCTCACCAGAGGTCCTACCAATCATATAGTGGAGGCACTGAAAGACCA 6708  
 Db 6301 TCTGACAGGCTCACCAGAGGTCCTACCAATCATATAGTGGAGGCACTGAAAGACCA 6360  
 QY 6709 GCGAGGCAATAGGTTGCGGAAGAGTGTGTTACCTGCGGCACTCTGTCAAGAGGCTT 6768  
 Db 6361 GCGAGGCAATAGGTTGCGGAAGAGTGTGTTACCTGCGGCACTCTGTCAAGAGGCTT 6420  
 QY 6769 GAACCAACCTACCGATGACTCTGCTGCTTGAACCCCTACACAGTCTCCCATTTATGCCCTGG 6828  
 Db 6421 GAACCAACCTACCGATGACTCTGCTGCTTGAACCCCTACACAGTCTCCCATTTATGCCCTGG 6480  
 QY 6829 AGATGAGTGGAACGAATGTTCTGAATCAGGCTTTAACTGTTGCTGCGAGTGTAGGCTT 6888  
 Db 6481 AGATGAGTGGAACGAATGTTCTGAATCAGGCTTTAACTGTTGCTGCGAGTGTAGGCTT 6540  
 QY 6889 TGGAGTGCTCATTTCCAGATGCTATCATCTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 6948  
 Db 6541 TGGAGTGCTCATTTCCAGATGCTATCATCTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 6600  
 QY 6949 CAAGATTTGAGAGAGAGTGGGACCGTCAAGGAGAAATGCGCAGATGATGAGCTGCAATG 7008  
 Db 6601 CAAGATTTGAGAGAGAGTGGGACCGTCAAGGAGAAATGCGCAGATGATGAGCTGCAATG 6660  
 QY 7009 TCTTGGAAACGGAAGAGAAATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7068  
 Db 6661 TCTTGGAAACGGAAGAGAAATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6720  
 QY 7069 TGGGAAGACATACCCAGTGGAGAACAGTGGGAGAAATATCTCGGTGCCATTTGCTC 7128  
 Db 6721 TGGGAAGACATACCCAGTGGAGAACAGTGGGAGAAATATCTCGGTGCCATTTGCTC 6780  
 QY 7129 CTGCAATGCTTTGGAGCCAGCGGGCTGCGCTGTGCAACTGCGCAGAGACTGCGGG 7188  
 Db 6781 CTGCAATGCTTTGGAGCCAGCGGGCTGCGCTGTGCAACTGCGCAGAGACTGCGGG 6840  
 QY 7189 TGAACCCAGTCCCGAAGGCACTACTGCGCCAGTCTCTCAACAGTATTCTCAGAGATACCA 7248  
 Db 6841 TGAACCCAGTCCCGAAGGCACTACTGCGCCAGTCTCTCAACAGTATTCTCAGAGATACCA 6900  
 QY 7249 TCAGAGAACAAACACTAATGTTTAAATGCTCCCAATGAGTGTCTTCAATCCAGAGAACCA 7308  
 Db 6901 TCAGAGAACAAACACTAATGTTTAAATGCTCCCAATGAGTGTCTTCAATCCAGAGAACCA 6960  
 QY 7309 GGTCTGACAGAGAGAAATCCCGAGAGTAAATCACTTTTCCAATCCAGAGAACCAAGCA 7368  
 Db 6961 GGTCTGACAGAGAGAAATCCCGAGAGTAAATCACTTTTCCAATCCAGAGAACCAAGCA 7020  
 QY 7369 CTCTCTGCCAAGATCTCAATCTAACTGGAGTGTGTTAGAGACCCAGCTTAGAGTTCTTC 7428



Db 661 CAAGAAGGATATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 720  
 QY 1069 GTGGAAGTGTGAGAGGACACCTCTCTGTGAGACCAACATCGAGCGGATCTGGCCCTTTCAC 1128  
 Db 721 GTGGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCGAGCGGATCTGGCCCTTTCAC 780  
 QY 1129 CGATGTTCTGTCAGCTGTTTACCAACCGGAGCCTCACCCCGGAGCTCTCCCTATGGGCA 1188  
 Db 781 CGATGTTCTGTCAGCTGTTTACCAACCGGAGCCTCACCCCGGAGCTCTCCCTATGGGCA 840  
 QY 1189 CTGTGTACAGACAGTGTGTGTCTTACTCTGTGGGATGCACTGGCTGAAGACACAAG 1248  
 Db 841 CTGTGTACAGACAGTGTGTGTCTTACTCTGTGGGATGCACTGGCTGAAGACACAAG 900  
 QY 1249 AATATGCAATCTTGTGACGTCCTGGGCAACGAGTCAGTCCGACAGACAGCTGT 1308  
 Db 901 AATATGCAATCTTGTGACGTCCTGGGCAACGAGTCAGTCCGACAGACAGCTGT 960  
 QY 1309 AACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTCTTACCATTCACTTACAA 1368  
 Db 961 AACCCAGACTTACGGTGGCAACTTAAATGGAGAGCCATGTCTTACCATTCACTTACAA 1020  
 QY 1369 TGGCAGACGTTTACTCTCTGCAACAGAGGCGACAGAGCAGACATCTTTGTGGAG 1428  
 Db 1021 TGGCAGACGTTTACTCTCTGCAACAGAGGCGACAGAGCAGACATCTTTGTGGAG 1080  
 QY 1429 CACAACTTCGAATATGAGCAGACAGAAATCTCTTCTGACAGACCACTTGTGTTT 1488  
 Db 1081 CACAACTTCGAATATGAGCAGACAGAAATCTCTTCTGACAGACCACTTGTGTTT 1140  
 QY 1489 GGTTCAGACTCGAGGAGAAATCCAAATGGTGTGCTTGTGCCATCTCCCTTCTCTATACAA 1548  
 Db 1141 GGTTCAGACTCGAGGAGAAATCCAAATGGTGTGCTTGTGCCATCTCCCTTCTCTATACAA 1200  
 QY 1549 CAACCAATTTACACTGATTTGCACTTCTGAGGCGACAGAGCAACATCAAGTGTGTGG 1608  
 Db 1201 CAACCAATTTACACTGATTTGCACTTCTGAGGCGACAGAGCAACATCAAGTGTGTGG 1260  
 QY 1609 GACCAACAGAACTATGATSCGACAGAAATTTGGTGTCTGCCCATCGCTGCCACGA 1668  
 Db 1261 GACCAACAGAACTATGATSCGACAGAAATTTGGTGTCTGCCCATCGCTGCCACGA 1320  
 QY 1669 GGAATCTGACACCAACCAATGAAGGCTCATGTACCGATTTGGAGTCAGTGGGATAGCA 1728  
 Db 1321 GGAATCTGACACCAACCAATGAAGGCTCATGTACCGATTTGGAGTCAGTGGGATAGCA 1380  
 QY 1729 GCATGACATGGGTACATGATGAGTGCACGCTGTGTGGGAATGTCGTGGGGAATGGAC 1788  
 Db 1381 GCATGACATGGGTACATGATGAGTGCACGCTGTGTGGGAATGTCGTGGGGAATGGAC 1440  
 QY 1789 ATGATTCGCTACTCGCAGCTTCAGATTCAGTGCATTTGTATGATCATCATTCATATGT 1848  
 Db 1441 ATGATTCGCTACTCGCAGCTTCAGATTCAGTGCATTTGTATGATCATCATTCATATGT 1500  
 QY 1849 GAACGACACATTCACAAGCGTTCATGAAGGCGCACATGCTGAACTGTACATGCTTCGG 1908  
 Db 1501 GAACGACACATTCACAAGCGTTCATGAAGGCGCACATGCTGAACTGTACATGCTTCGG 1560  
 QY 1909 TCAGGGTCGGGCGAGTGTGAATGTCCTCGACCAATGCGCAATTCAGAGACTGG 1968  
 Db 1561 TCAGGGTCGGGCGAGTGTGAATGTCCTCGACCAATGCGCAATTCAGAGACTGG 1620  
 QY 1969 GACGTTTTATCAATTTGAGATTCATGGGAGAGTATGTGATGTTGATGATACAGTCCAGT 2028  
 Db 1621 GACGTTTTATCAATTTGAGATTCATGGGAGAGTATGTGATGTTGATGATACAGTCCAGT 1680  
 QY 2029 CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCCAATTTTACAGACCTTATCCAG 2088  
 Db 1681 CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCCAATTTTACAGACCTTATCCAG 1740  
 QY 2089 CTCAAGTGGTCTGTGCAAGTATTTATCACTGATCTCCGAGTCAGCCCACTCCACCC 2148

Db 1741 CTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAGCCCACTCCACCC 1800  
 QY 2149 CATCCAGTGAATGCAACCAAGCACTCTCACTTTCAAGTACATTTCTAGTGGAGACC 2208  
 Db 1801 CATCCAGTGAATGCAACCAAGCACTCTCACTTTCAAGTACATTTCTAGTGGAGACC 1860  
 QY 2209 TAAAAATTTCTGAGGCGGTTGGAGGAGTCAATACACAGGCGCACTTAAATCTCTACAC 2268  
 Db 1861 TAAAAATTTCTGAGGCGGTTGGAGGAGTCAATACACAGGCGCACTTAAATCTCTACAC 1920  
 QY 2269 CATCAAGGCTGAAGCCTGTGTGTATACAGGCGCAGCTCATCAGCATCCAGAGTA 2328  
 Db 1921 CATCAAGGCTGAAGCCTGTGTGTATACAGGCGCAGCTCATCAGCATCCAGAGTA 1980  
 QY 2329 CGGCCACCAAGAGTACTCGCTTTGACTTACACACAGCAGCAGCAGCAGCTGTGAC 2388  
 Db 1981 CGGCCACCAAGAGTACTCGCTTTGACTTACACACAGCAGCAGCAGCAGCTGTGAC 2040  
 QY 2389 CAGCAACACCGTGAAGAGAGAGAGCTCCCTTTCTCTTGTGGCCACTTCTGAATC 2448  
 Db 2041 CAGCAACACCGTGAAGAGAGAGAGCTCCCTTTCTCTTGTGGCCACTTCTGAATC 2100  
 QY 2449 TGTGACCGAATCACAGCCAGTACTTGTGTCTCTCTGGTCTCAGCTTCCGACACCGT 2508  
 Db 2101 TGTGACCGAATCACAGCCAGTACTTGTGTCTCTCTGGTCTCAGCTTCCGACACCGT 2160  
 QY 2509 GTCGGGATTCGGGTGGAATATGAGCTGAGTGGAGGAGATGAGCCACAGTACCTTGA 2568  
 Db 2161 GTCGGGATTCGGGTGGAATATGAGCTGAGTGGAGGAGATGAGCCACAGTACCTTGA 2220  
 QY 2569 TCTTCCAGACAGCCACTTCTGTGAATCTCTGACCTGCTTCTGGCGGAAATACAT 2628  
 Db 2221 TCTTCCAGACAGCCACTTCTGTGAATCTCTGACCTGCTTCTGGCGGAAATACAT 2280  
 QY 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTGTCTACTTCA 2688  
 Db 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGAGTTTGTATCTGTCTACTTCA 2340  
 QY 2889 AACAAAGCGCTGATCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2748  
 Db 2341 AACAAAGCGCTGATCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2400  
 QY 2749 TGTGTTGCTGGAGCAGACCCAGGCTCCATCAGAGGTACAGAAATAGTCTATTGCC 2808  
 Db 2401 TGTGTTGCTGGAGCAGACCCAGGCTCCATCAGAGGTACAGAAATAGTCTATTGCC 2460  
 QY 2809 ATCAGTGAAGGTAGCAGCAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCT 2868  
 Db 2461 ATCAGTGAAGGTAGCAGCAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCT 2520  
 QY 2869 CAGTGACTTGCAACCTGCTGATACATCACTATCTATGCTGTGGAGAAATATCA 2928  
 Db 2521 CAGTGACTTGCAACCTGCTGATACATCACTATCTATGCTGTGGAGAAATATCA 2580  
 QY 2929 AGAAAGTACACCTTGTCTATTCACCAAGAAACCACTGGCACCCCGCTCAGATACAGT 2988  
 Db 2591 AGAAAGTACACCTTGTCTATTCACCAAGAAACCACTGGCACCCCGCTCAGATACAGT 2640  
 QY 2989 GCGCTCTCCAGGAGCCTGCACTTGTGGAGTACAGAGCTGAGGTCACCTCATGTG 3048  
 Db 2641 GCGCTCTCCAGGAGCCTGCACTTGTGGAGTACAGAGCTGAGGTCACCTCATGTG 2700  
 QY 3049 GACACCGCTCAGAGTGCAGTGCAGGCTACCGGTGTGGATGTGATCCCGCTCAACCTGCC 3108  
 Db 2701 GACACCGCTCAGAGTGCAGTGCAGGCTACCGGTGTGGATGTGATCCCGCTCAACCTGCC 2760  
 QY 3109 TGGCGAGCAGGCGAGGCTGCCCATCAGCAGGAAACCTTTTGCAGAGTCCAGGGCT 3168  
 Db 2761 TGGCGAGCAGGCGAGGCTGCCCATCAGCAGGAAACCTTTTGCAGAGTCCAGGGCT 2820  
 QY 3169 GTCCCTTGGGTCACCTATTACTTCAAAAGTCTTTGAGTGCAGCTGAGGAGGAGAGCAA 3228  
 Db 2821 GTCCCTTGGGTCACCTATTACTTCAAAAGTCTTTGAGTGCAGCTGAGGAGGAGAGCAA 2880





Db 5041 TGAATGAGAGCCAGCCCTGATTTGAAACCCAGTCCACAGATATTTCTGACCAACTGA 5100  
Qy 5449 CTTGAAGTTCACTCAGGTCAACCCACAAAGCTGAGCGCCAGTGAGACACCAACCAATGT 5508  
Db 5101 CTTGAAGTTCACTCAGGTCAACCCACAAAGCTGAGCGCCAGTGAGACACCAACCAATGT 5160  
Qy 5509 TGAAGTCACTGATGATGAGTGGGTGACCCCAAGAGAGAGACCGGACCAATGAAGA 5568  
Db 5161 TGAAGTCACTGATGATGAGTGGGTGACCCCAAGAGAGAGACCGGACCAATGAAGA 5220  
Qy 5569 AATCAACCTTGTCTGACAGCTCACTCGGTGTTGATCAGAGCTATGTTGGTGGCCACCAA 5628  
Db 5221 AATCAACCTTGTCTGACAGCTCACTCGGTGTTGATCAGAGCTATGTTGGTGGCCACCAA 5280  
Qy 5629 ATATGAAGTGAAGTGTATGCTCTTAAGGACACTTTTGAACAAGACACAGCTCAGGGAGT 5688  
Db 5281 ATATGAAGTGAAGTGTATGCTCTTAAGGACACTTTTGAACAAGACACAGCTCAGGGAGT 5340  
Qy 5689 TGTCAACCACTCTGAGAGATGTCAGCCCAACAGAGGGCTCGTGTGACAGATGCTACTGA 5748  
Db 5341 TGTCAACCACTCTGAGAGATGTCAGCCCAACAGAGGGCTCGTGTGACAGATGCTACTGA 5400  
Qy 5749 GACCACCACTCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCCCTCCAAAGTTGA 5808  
Db 5401 GACCACCACTCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGA 5460  
Qy 5809 TGGCGTTCAGCCCAATGGCCAGACTCCATCCAGAGAACCATCAAGCCAGATGTCAGAG 5868  
Db 5461 TGGCGTTCAGCCCAATGGCCAGACTCCATCCAGAGAACCATCAAGCCAGATGTCAGAG 5520  
Qy 5869 CTACACCACTCACTGGCTTCAACCAAGCCAGCTGACTACAAGATCTACTGTACACCTTGAA 5928  
Db 5521 CTACACCACTCACTGGCTTCAACCAAGCCAGCTGACTACAAGATCTACTGTACACCTTGAA 5580  
Qy 5929 TGAATGCTCGAGCTCCCTGTGTATGAGAGCCCTGAGTCTCTCCAGAGAGAT 6108  
Db 5581 TGAATGCTCGAGCTCCCTGTGTATGAGAGCCCTGAGTCTCTCCAGAGAGAT 5760  
Qy 5989 CAACCTTGGCTTCTGGCCACCAACCCCAATCTCTGCTGATCATGCGAGCCGACG 6048  
Db 5641 CAACCTTGGCTTCTGGCCACCAACCCCAATCTCTGCTGATCATGCGAGCCGACG 5700  
Qy 6049 TGGCAGATTAACGGCTTACATCAATCAAGTATGAGAGCCCTGAGTCTCTCCAGAGAGAT 6108  
Db 5701 TGGCAGATTAACGGCTTACATCAATCAAGTATGAGAGCCCTGAGTCTCTCCAGAGAGAT 5760  
Qy 6109 GGTCCCTCGCCCGCCCTGCTGTCAGAGGCTACTATTACTGGCTTGAACCCGGAAC 6168  
Db 5761 GGTCCCTCGCCCGCCCTGCTGTCAGAGGCTACTATTACTGGCTTGAACCCGGAAC 5820  
Qy 6169 CGAATATACAAATTAATGCTATGCTTCCCTGAAAGATAATCAAGAGAGCGCCCTGATGG 6228  
Db 5821 CGAATATACAAATTAATGCTATGCTTCCCTGAAAGATAATCAAGAGAGCGCCCTGATGG 5880  
Qy 6229 AAGGAAAGACAGACAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTCATGG 6288  
Db 5881 AAGGAAAGACAGACAGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTCATGG 5940  
Qy 6289 ACCAGAGATCTTGGATGCTTCTCCACAGTTCAAAAGACCCCTTCTGTCACCCACCCCTGG 6348  
Db 5941 ACCAGAGATCTTGGATGCTTCTCCACAGTTCAAAAGACCCCTTCTGTCACCCACCCCTGG 6000  
Qy 6349 GTATGACACTGGAAATGGTATGAGTCTCTGCACTTCTGTCAGCAACCCAGTGTGG 6408  
Db 6001 GTATGACACTGGAAATGGTATGAGTCTCTGCACTTCTGTCAGCAACCCAGTGTGG 6060  
Qy 6409 GCAACAAATGATCTTTCAGAGAACATGTTTATGGCGGACCAACCCGACCAACAGCCGAC 6468  
Db 6061 GCAACAAATGATCTTTCAGAGAACATGTTTATGGCGGACCAACCCGACCAACAGCCGAC 6120  
Qy 6469 CCCCATAGGCATAGGCCAAGCAACATACCCGCGAATGTAGGACAAAGAGCTCTCTCTCA 6528

Db 6121 CCCCATAGGCATAGGCCAAGACCATACCCGCGAATGTAGGACAAAGAGCTCTCTCTCA 6180  
Qy 6529 GACAAACCATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCATTTTCAATGTCATCC 6588  
Db 6181 GACAAACCATCTCATGGGCCCATTCAGGACACTTCTGAGTACATCATTTTCAATGTCATCC 6240  
Qy 6589 TGTGGCACATGATGAAGAACCCCTTACAGTTTCAAGGTTCTTGGAACTTCTTACCAAGTCCAC 6648  
Db 6241 TGTGGCACATGATGAAGAACCCCTTACAGTTTCAAGGTTCTTGGAACTTCTTACCAAGTCCAC 6300  
Qy 6649 TCTGACAGGCTTCAACAGAGGTCCTCAACATCATAGTGGAGGCACTGAAAGACCA 6708  
Db 6301 TCTGACAGGCTTCAACAGAGGTCCTCAACATCATAGTGGAGGCACTGAAAGACCA 6360  
Qy 6709 GCAGAGCATTAAGTTCGGGAAGAGGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6768  
Db 6361 GCAGAGCATTAAGTTCGGGAAGAGGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6420  
Qy 6769 GAACCAACCTACGAGTGACTGTGCTTTGACCCCTACACAGTGTCCCATTTATGCCGTGG 6828  
Db 6421 GAACCAACCTACGAGTGACTGTGCTTTGACCCCTACACAGTGTCCCATTTATGCCGTGG 6480  
Qy 6829 AGATGAGTGGGAAGATGTCTGAATCAGGCTTTAACTGTTGTCAGGCTTAAAGCTT 6888  
Db 6481 AGATGAGTGGGAAGATGTCTGAATCAGGCTTTAACTGTTGTCAGGCTTAAAGCTT 6540  
Qy 6889 TGGAAAGTGGTCAATTTCAAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAAC 6948  
Db 6541 TGGAAAGTGGTCAATTTCAAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAAC 6600  
Qy 6949 CAAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGTGCACATG 7008  
Db 6601 CAAGATTTGGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGTGCACATG 6660  
Qy 7009 TCTTGGGAAACGGAAAGAGAAATTCAGTGTGACCTCATGAGGCAACGTTTACGATGA 7068  
Db 6661 TCTTGGGAAACGGAAAGAGAAATTCAGTGTGACCTCATGAGGCAACGTTTACGATGA 6720  
Qy 7069 TGGAAAGACATPACCACTGAGAGAACAGTGGCAGAAATATCTCGGTGCCATTTGCTC 7128  
Db 6721 TGGAAAGACATPACCACTGAGAGAACAGTGGCAGAAATATCTCGGTGCCATTTGCTC 6780  
Qy 7129 CTGCACATGCTTTGGAGCCAGCGGGCTGCGCGTGTGACAACTGCGCGAGACCTTGGGG 7188  
Db 6781 CTGCACATGCTTTGGAGCCAGCGGGCTGCGCGTGTGACAACTGCGCGAGACCTTGGGG 6840  
Qy 7189 TGAACCCAGTCCCGAAGGCACTACTGCGCCAGTCTTACAAACAGTATTTCTCAGAGATACCA 7248  
Db 6841 TGAACCCAGTCCCGAAGGCACTACTGCGCCAGTCTTACAAACAGTATTTCTCAGAGATACCA 6900  
Qy 7249 TCAGAGAACAAACACTAATGTTAATTCGCCAATGAGTGTGCTTCATGCTTTAGATGTACA 7308  
Db 6901 TCAGAGAACAAACACTAATGTTAATTCGCCAATGAGTGTGCTTCATGCTTTAGATGTACA 6960  
Qy 7309 GCCTGACAGAGAGATTTCCCGAGAGTAAATCATCTTTTCCAATCCAGAGGAACCAAGCATGT 7368  
Db 6961 GCCTGACAGAGAGATTTCCCGAGAGTAAATCATCTTTTCCAATCCAGAGGAACCAAGCATGT 7020  
Qy 7369 CTCTCTGCAAGATGCATCTTAACTGAGTGTATGTTAGCAGACCCAGCTTAGAGTCTTCT 7428  
Db 7021 CTCTCTGCAAGATGCATCTTAACTGAGTGTATGTTAGCAGACCCAGCTTAGAGTCTTCT 7080  
Qy 7429 TTTCTTTCTTAAAGCCCTTGTCTCTGAGGAGATTTCTCCAGCTTCCAGCTCAACTCAGAGCT 7488  
Db 7081 TTTCTTTCTTAAAGCCCTTGTCTCTGAGGAGATTTCTCCAGCTTCCAGCTCAACTCAGAGCT 7140  
Qy 7489 TCTTCAAGCATCAACCTGGAGTTCCTGAGGTTTCTGAGGTTTCTCATAAATGAGGGCTGCACAT 7548  
Db 7141 TCTTCAAGCATCAACCTGGAGTTCCTGAGGTTTCTGAGGTTTCTCATAAATGAGGGCTGCACAT 7200  
Qy 7549 GCCTGTTCTGCTTCCGAGATTTCAATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7608  
Db 7201 GCCTGTTCTGCTTCCGAGTATTTCAATACCGCTCAGTATTTTAAATGAAGTGAATCTAAGA 7260

2y 7609 TTTCGTTGGATCAATAGGAAGCATATGCGAGCAACCAAGATGCAATGTTTGAAT 7668  
 2b 7261 TTTCGTTGGATCAATAGGAAGCATATGCGAGCAACCAAGATGCAATGTTTGAAT 7320  
 2y 7669 GATATGACCAAAATTTTAAGTAGGAAGTCAACCAACACTTCTGCTTTTCACTTAAGTGT 7728  
 2b 7321 GATATGACCAAAATTTTAAGTAGGAAGTCAACCAACACTTCTGCTTTTCACTTAAGTGT 7380  
 2y 7729 CTGCGCCGCAATCTGTAGGAACAAGCATGATCTGTGATCTGTGATATTTAAATATCCA 7788  
 2b 7381 CTGCGCCGCAATCTGTAGGAACAAGCATGATCTGTGATATTTAAATATCCA 7440  
 2y 7789 CAGTACT 7795  
 2b 7441 CAGTACT 7447

RESULT 14  
 ACF12859  
 ID ACF12859 standard; cDNA; 7680 BP.  
 KC ACF12859;  
 KC 10-SEP-2003 (first entry)  
 JT Human cervical cancer cell marker encoding cDNA SEQ ID NO:63.  
 DE Human; cervical cancer; cervical cancer marker; cancer therapy;  
 CW detection; gene therapy; vaccine; gene; ss.  
 KW Homo sapiens.  
 JS WO2002101075-A2.  
 PN 19-DEC-2002.  
 2D 12-JUN-2002; 2002MO-US018638.  
 2K 13-JUN-2001; 2001US-0298155P.  
 PR 13-JUN-2001; 2001US-0298159P.  
 PR 14-NOV-2001; 2001US-0335936P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 FA Schlegel R, Chen Y, Zhao X, Monahan JB, Kamatkar S;  
 PI Gannavarapu M, Glatt K, Hoersch S;  
 XX WPI; 2003-155967/15.  
 DR P-PSDB; ABR92078.  
 XX New isolated nucleic acid molecule useful for detecting, characterizing,  
 PT preventing and treating human cervical cancers, in various prognostic and  
 PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.  
 XX Claim 4; Page 209-212; 386pp; English.

ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)  
 given in ABR92047 to ABR92164. A higher level of expression of (I) than  
 normal indicates the presence of cervical cancer. Also described: (1) a  
 vector (ii) containing (I); (2) a host cell (iii) containing (I); and (3)  
 assessing (M) whether a patient is afflicted with cervical cancer.  
 CC comprising comparing the level of expression of a marker in a patient's  
 sample, and the normal level of expression of the marker in a control non  
 CC cervical cancer sample, where a significant increase in the level of  
 CC expression of the marker in the patient's sample relative to that in the  
 CC control sample is an indication that the patient is afflicted with  
 CC cervical cancer. (I) has cytostatic activity, and can be used in gene  
 CC therapy and in vaccines. (I) is useful in detecting, characterizing,  
 CC preventing and treating human cervical cancers. (I) may also be used in  
 CC various prognostic and diagnostic assays, pharmacogenomics and in  
 CC monitoring clinical trials

SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;  
 Query Match 95.2%; Score 7423; DB 7; Length 7680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 349 GAAGAGCAAGAGGAGGCTCAGCAAAATGTTTTCAGCCCGAGTCCCGGGTGGCTGTCACTCA 408  
 Db 1 GAAGAGCAAGAGGAGGCTCAGCAAAATGTTTTCAGCCCGAGTCCCGGGTGGCTGTCACTCA 60  
 QY 409 AAGCAAGCCCGGTTTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCG 468  
 Db 61 AAGCAAGCCCGGTTTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCG 120  
 QY 469 GACCTACCTAGGCAATGCTTGTGTTTGTACCTTGTATGGAGGAGCCGAGGTTTAACTG 528  
 Db 121 GACCTACCTAGGCAATGCTTGTGTTTGTACCTTGTATGGAGGAGCCGAGGTTTAACTG 180  
 QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGAACAGTACACTGGGAACACTTACCG 588  
 Db 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGAACAGTACACTGGGAACACTTACCG 240  
 QY 589 AGTGGGTGACACTTATGAGCGCTCTTAAGACTCCATGATCTGGGACTGTACTTCGATCGG 648  
 Db 241 AGTGGGTGACACTTATGAGCGCTCTTAAGACTCCATGATCTGGGACTGTACTTCGATCGG 300  
 QY 649 GGCTGGGCGAGGAGATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCACTC 708  
 Db 301 GGCTGGGCGAGGAGATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCACTC 360  
 QY 709 CTACAAGATTGGTGACACTCTGGAGGAGACCACTGAGACTGGTGTACATGTTAGAGTG 768  
 Db 361 CTACAAGATTGGTGACACTCTGGAGGAGACCACTGAGACTGGTGTACATGTTAGAGTG 420  
 QY 769 TGTGTGTCTTGGTATGAAAAGGAGAAATGACCTGCAAGCCCATAGCTGAGAAAGTGT 828  
 Db 421 TGTGTGTCTTGGTATGAAAAGGAGAAATGACCTGCAAGCCCATAGCTGAGAAAGTGT 480  
 QY 829 TGATCATGCTCTGGGACTTCTATGTGTGCGAGAAACGTGGGAGAAAGCCCTTACCAAGG 888  
 Db 481 TGATCATGCTCTGGGACTTCTATGTGTGCGAGAAACGTGGGAGAAAGCCCTTACCAAGG 540  
 QY 889 CTGGATGATGTTAGATTGTACTTGGCTGGGAGAGGCGGACGACATCACTTGGACTTC 948  
 Db 541 CTGGATGATGTTAGATTGTACTTGGCTGGGAGAGGAGCGGACGACATCACTTGGACTTC 600  
 QY 949 TAGAAATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATTGGAGACACTGGAG 1008  
 Db 601 TAGAAATAGATGCAACGATCAGGACACAAGGACATCTTATAGAAATTGGAGACACTGGAG 660  
 QY 1009 CAAGAGGATATCGAGGAAACCTCTCTCAGTGCATCTGCACAGCAACGCGCGAGGAGA 1068  
 Db 661 CAAGAGGATATCGAGGAAACCTCTCTCAGTGCATCTGCACAGCAACGCGCGAGGAGA 720  
 QY 1069 GTGGAAGTGTGAGAGGCGACACCTCTGTGCAGACCACTCGAGCGGATCTGGCCCTTCA 1128  
 Db 721 GTGGAAGTGTGAGAGGCGACACCTCTGTGCAGACCACTCGAGCGGATCTGGCCCTTCA 780  
 QY 1129 CGATGTTGTGAGCTGTTTACCAACCGAGCTCACCCAGCCCTCTCCCTATGGCCA 1188  
 Db 781 CGATGTTGTGAGCTGTTTACCAACCGAGCTCACCCAGCCCTCTCCCTATGGCCA 840  
 QY 1189 CTGTGTACAGACAGTGTGTGTGTCTCTGTGGGATGCGAGTGGCTGGAAGACACAAGG 1248  
 Db 841 CTGTGTACAGACAGTGTGTGTGTCTCTGTGGGATGCGAGTGGCTGGAAGACACAAGG 900  
 QY 1249 AAATAGCAAAATGTTTGCATGCTGGGCAACGAGTCACTGCCAAGAGACAGCTGT 1308  
 Db 901 AAATAGCAAAATGTTTGCATGCTGGGCAACGAGTCACTGCCAAGAGACAGCTGT 960  
 QY 1309 AACCAGAGCTTACGTTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA 1368  
 Db 961 AACCAGAGCTTACGTTGGCAACTTAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA 1020

QY 1369 TGGCAGGAGCTTCTACTCTCGACCAAGAGGCGACAGACGAGACATCTTTGTGTCAG 1428  
Db 1021 TGGCAGGAGCTTCTACTCTCTCGACCAAGAGGCGACAGACGAGACATCTTTGTGTCAG 1080  
QY 1429 CACAACTTCGAATTTATGAGCAGGACAGAAATACTCTTTCTGACAGACCAACACATGTTT 1488  
Db 1081 CACAACTTCGAATTTATGAGCAGGACAGAAATACTCTTTCTGACAGACCAACACATGTTT 1140  
QY 1489 GGTTCAGACTCGAGGAGGAAATTCCTAATGGTGGCTTGTGCCACTTCCCTTCTCTATACAA 1548  
Db 1141 GGTTCAGACTCGAGGAGGAAATTCCTAATGGTGGCTTGTGCCACTTCCCTTCTCTATACAA 1200  
QY 1549 CAACCAATTTACACTGATGACATCTCTGAGGCGAGAGACACATGAGTGTGTTGG 1608  
Db 1201 CAACCAATTTACACTGATGACATCTCTGAGGCGAGAGACACATGAGTGTGTTGG 1260  
QY 1609 GACCAACAGAACTATGATGCCAGCAGAAATTTGGTTCCTCCCAATGGCTGCCACGA 1668  
Db 1261 GACCAACAGAACTATGATGCCAGCAGAAATTTGGTTCCTCCCAATGGCTGCCACGA 1320  
QY 1669 GGAATCTGCACACCAATGAGGGGTCTATGTCGATGAGATGAGTGGGTAAGCA 1728  
Db 1321 GGAATCTGCACACCAATGAGGGGTCTATGTCGATGAGATGAGTGGGTAAGCA 1380  
QY 1729 GCATGATGGGTACATGATGAGGTGCACGTGTGTTGGGAATGGTGGGGAATGGAC 1788  
Db 1381 GCATGATGGGTACATGATGAGGTGCACGTGTGTTGGGAATGGTGGGGAATGGAC 1440  
QY 1789 ATGCATTCCTTCTCGAGCTTCGAGATCAGTGCATTTGTGATGATCACTTCAATGT 1848  
Db 1441 ATGCATTCCTTCTCGAGCTTCGAGATCAGTGCATTTGTGATGATCACTTCAATGT 1500  
QY 1849 GAACGACACATTCACAAAGCTTCATCAAGAGGGGCACATGCTGAACTGTACATGCTTGG 1908  
Db 1501 GAACGACACATTCACAAAGCTTCATCAAGAGGGGCACATGCTGAACTGTACATGCTTGG 1560  
QY 1909 TCAGGTGGGGCAGTGGAGTGTGATCCCGTCGACCAATGCCAGATTCAGAGACTGG 1968  
Db 1561 TCAGGTGGGGCAGTGGAGTGTGATCCCGTCGACCAATGCCAGATTCAGAGACTGG 1620  
QY 1969 GAGTTTTATCAATTTGGAGTTTCATGGGAGTATGTCATGTTGTCAGATACAGTGG 2028  
Db 1621 GAGTTTTATCAATTTGGAGTTTCATGGGAGTATGTCATGTTGTCAGATACAGTGG 1680  
QY 2029 CTACTGTATGGCGGTGGCATTCGGGAGTGGCATTCGCCAACCTTTACAGACCTATCCAA 2088  
Db 1681 CTACTGTATGGCGGTGGCATTCGGGAGTGGCATTCGCCAACCTTTACAGACCTATCCAA 1740  
QY 2089 CTCAAGTGTCTGTGAGATTTTATCACTGAGACTCCGATCGACCCCACTCCCAACC 2148  
Db 1741 CTCAAGTGTCTGTGAGATTTTATCACTGAGACTCCGATCGACCCCACTCCCAACC 1800  
QY 2149 CATCCAGTGAATGCACACAGCCATCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACC 2208  
Db 1801 CATCCAGTGAATGCACACAGCCATCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACC 1860  
QY 2209 TAAAAATTTGTAGGCGTTGGAGGAGTACCATACAGGCCACTTAACTCCTACAC 2268  
Db 1861 TAAAAATTTGTAGGCGTTGGAGGAGTACCATACAGGCCACTTAACTCCTACAC 1920  
QY 2269 CATCAAGGCTCAAGCCTGGTGTGTATACAGAGGCCAGCTCATCAGCATCCAGCAGTA 2328  
Db 1921 CATCAAGGCTCAAGCCTGGTGTGTATACAGAGGCCAGCTCATCAGCATCCAGCAGTA 1980  
QY 2329 CGGCCACCAAGAGTGAATGCTTGTGATTCACCAACACAGCAGCAGCACTGTGAC 2388  
Db 1981 CGGCCACCAAGAGTGAATGCTTGTGATTCACCAACACAGCAGCAGCACTGTGAC 2040  
QY 2389 CAGCAACACCGTGCAGGAGAGACTCCCTTTCTCTTGTGGCCACTTCTGAATC 2448  
Db 2041 CAGCAACACCGTGCAGGAGAGACTCCCTTTCTCTTGTGGCCACTTCTGAATC 2100

QY 2449 TGTACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2508  
Db 2101 TGTACCGAAATCACAGCCAGTAGCTTTGTGGTCTCTCTGGGTCTCAGCTTCCGACACCGT 2160  
QY 2509 GTCCGGATTCGGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2568  
Db 2161 GTCCGGATTCGGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2220  
QY 2569 TCTTCCAGCAGCAGCCACTTCTGTGAACATCCCTGACCTGCTTCTCTGGCGGAAAAATACAT 2628  
Db 2221 TCTTCCAGCAGCAGCCACTTCTGTGAACATCCCTGACCTGCTTCTCTGGCGGAAAAATACAT 2280  
QY 2629 TGTAAAATGCTCTATCAGATATCTCAGGATGGGAGCAGAGTTTGTATCTGTCTACTTCA 2688  
Db 2281 TGTAAAATGCTCTATCAGATATCTCAGGATGGGAGCAGAGTTTGTATCTGTCTACTTCA 2340  
QY 2689 AACAAAGCCCTGTATGTCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2748  
Db 2341 AACAAAGCCCTGTATGTCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2400  
QY 2749 TGTGTTTCGCTGGAGCAGACCCAGGCTCCCATCACAGGGTACAGAAATAGTCTATTGCGC 2808  
Db 2401 TGTGTTTCGCTGGAGCAGACCCAGGCTCCCATCACAGGGTACAGAAATAGTCTATTGCGC 2460  
QY 2809 ATCAGTAAAGGTAGCAGCAGAACTCAACCTTCTGAAAACCTGCAAACTCCCTCACCT 2868  
Db 2461 ATCAGTAAAGGTAGCAGCAGAACTCAACCTTCTGAAAACCTGCAAACTCCCTCACCT 2520  
QY 2869 CAGTGAATTCGAACCTGGTGTTCAGTATATACATCACTATCTATGCTGTGGAAGAAATCA 2928  
Db 2521 CAGTGAATTCGAACCTGGTGTTCAGTATATACATCACTATCTATGCTGTGGAAGAAATCA 2580  
QY 2929 AGAAAGTACACCTGTTGTCTATTCAACAAGAAACCTCTGGCACCCCACTCAGATACAGT 2988  
Db 2581 AGAAAGTACACCTGTTGTCTATTCAACAAGAAACCTCTGGCACCCCACTCAGATACAGT 2640  
QY 2989 GCCCTTCCAGGAGCCTGCACTTGTGGAGTGCAGAGCTGAAGTCACTCATGTG 3048  
Db 2641 GCCCTTCCAGGAGCCTGCACTTGTGGAGTGCAGAGCTGAAGTCACTCATGTG 2700  
QY 3049 GACACCCCTTGAAGTGCAGTGCAGGCTACCGTGTGGATGTGATCCCCCTCAACCTGCC 3108  
Db 2701 GACACCCCTTGAAGTGCAGTGCAGGCTACCGTGTGGATGTGATCCCCCTCAACCTGCC 2760  
QY 3109 TGGCGAGCAGCGGAGGCTGCCATCAGCAGAAACCTTTTGCAGAACTCAGCGGCT 3168  
Db 2761 TGGCGAGCAGCGGAGGCTGCCATCAGCAGAAACCTTTTGCAGAACTCAGCGGCT 2820  
QY 3169 GTCCCTTGGGCTCACCTATTACTTCAAGTCTTTGAGTGCAGTGCAGGAGGAGAGCAA 3228  
Db 2821 GTCCCTTGGGCTCACCTATTACTTCAAGTCTTTGAGTGCAGTGCAGGAGGAGAGCAA 2880  
QY 3229 GCCTCTGACTGCTCAACAGACAAACCTGATGCTTCCCACTAACTCCAGTTTGTCAA 3288  
Db 2881 GCCTCTGACTGCTCAACAGACAAACCTGATGCTTCCCACTAACTCCAGTTTGTCAA 2940  
QY 3289 TGAACCTGATTTCTCTCTGCTGATGATGCTCCACTCCGCGGCCAGATAACAGATA 3348  
Db 2941 TGAACCTGATTTCTCTCTGCTGATGATGCTCCACTCCGCGGCCAGATAACAGATA 3000  
QY 3349 CCGACTGACCGTGGGCTTACCCGAGAGGACAGCCAGGAGTACATGTTGGGTCCCTC 3408  
Db 3001 CCGACTGACCGTGGGCTTACCCGAGAGGACAGCCAGGAGTACATGTTGGGTCCCTC 3060  
QY 3409 TGTCTCAAGTACCCATGAGGATCTGACGCTGATCTGAGTACACCTATCCCTCGT 3468  
Db 3061 TGTCTCAAGTACCCCTGAGGATCTGACGCTGATCTGAGTACACCTATCCCTCGT 3120  
QY 3469 GGCCATAAAGGGCAACCAAGAGAGCCCAAGGAGCTTGAAGTCTTTACCACTGCGAGCC 3528  
Db 3121 GGCCATAAAGGGCAACCAAGAGAGCCCAAGGAGCTTGAAGTCTTTACCACTGCGAGCC 3180  
QY 3529 TGGGAGCTCTATTCCACTTACACCCGAGGTGACTGAGACCACTTGTGATGATG 3588

[illegible]

4261 TTCTGATGTTCCGAGGACCTGGAAAGTTGTGTGCTGGACACCCCCACACAGCCTACTGATCAG 4320

4669 CTGGGATGCTCCTGCTGTCA CAGTGAGATATTA CAGGATCACTTTACGGAGAAA CAGGAGG 4728

4321 CTGGGATGCTCCTGCTGTCA CAGTGAGATATTA CAGGATCACTTTACGGAGAAA CAGGAGG 4380

4729 AAATAGCCCTGTCCAGAGTTCACTGTGCTCGGAGCAAGTCTACAGCTACCATCAGCGG 4788

4381 AAATAGCCCTGTCCAGAGTTCACTGTGCTCGGAGCAAGTCTACAGCTACCATCAGCGG 4440

4789 CTTTAAACCTGGAGTGAATATACCAATCACTGTGATGCTGTCACTGGCCGTGGAGACAG 4848

4441 CTTTAAACCTGGAGTGAATATACCAATCACTGTGATGCTGTCACTGGCCGTGGAGACAG 4500

4849 CCCCGAAGCAGCAGCCAAATTTCCAATTAATTCAGAA CAGAAATAGCAAAACCATCCCA 4908

4501 CCCCGAAGCAGCAGCCAAATTTCCAATTAATTAACCGAA CAGAAATAGCAAAACCATCCCA 4560

4909 GATGCAAGTGACCGATGTT CAGGACAA CAGCAATTAGTGTCAAGTGGCTGCCTTCAAGTTC 4968

4561 GATGCAAGTGACCGATGTT CAGGACAA CAGCAATTAGTGTCAAGTGGCTGCCTTCAAGTTC 4620

4969 CCCTGTTACTGGTTTACAGAGTAACACCACTCCCAAAAATGGACAGGACCAACAAAAC 5028

4621 CCCTGTTACTGGTTTACAGAGTAACACCACTCCCAAAAATGGACAGGACCAACAAAAC 4680

5029 TAAACCTGCAGTCCAGATCAACACAGAAATGACTATTGCAAGCTTGCAGCCACAGTGGGA 5088

4681 TAAACCTGCAGTCCAGATCAACACAGAAATGACTATTGCAAGCTTGCAGCCACAGTGGGA 4740

5089 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGTCAAGCTCTGGTTCAGAC 5148

4741 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGTCAAGCTCTGGTTCAGAC 4800

5149 TGCAGTAACCAACATTTGATCGCCCTTAAAGGACTGGCATTTCACTGATGGATGTCGATTC 5208

4801 TGCAGTAACCAACATTTGATCGCCCTTAAAGGACTGGCATTTCACTGATGGATGTCGATTC 4860

5209 CATCAAAATTTGCTTGGAAAGCCCA CAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 5268

4861 CATCAAAATTTGCTTGGAAAGCCCA CAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 4920

5269 GAGCCCTGAGATGGAAATCAATGAGCTATTCCCTGCACTGATGGTGAAGAGACACTGC 5328

4921 GAGCCCTGAGATGGAAATCAATGAGCTATTCCCTGCACTGATGGTGAAGAGACACTGC 4980

5329 AGAGCTGCAAGGCTCAGACCGGGTTCTCAGATPACAGTCAAGTGTGGTGTGCTTGCACGA 5388

4981 AGAGCTGCAAGGCTCAGACCGGGTTCTCAGATPACAGTCAAGTGTGGTGTGCTTGCACGA 5040

5389 TGATATGGAGAGCAGCCCTGATTTGGAAACCAAGTCCA CAGCTATTCTCTGCAACCACTGA 5448

5041 TGATATGGAGAGCAGCCCTGATTTGGAAACCAAGTCCA CAGCTATTCTCTGCAACCACTGA 5100

5449 CCTGAGTTCATCTCAGGTCAACCCACAAAGCTGAGCGCCAGTGGACACCAACCCCAATGT 5508

5101 CCTGAGTTCATCTCAGGTCAACCCACAAAGCTGAGCGCCAGTGGACACCAACCCCAATGT 5160

5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAAACCCGAGACCAATGAAAGA 5568

5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAAACCCGAGACCAATGAAAGA 5220

5569 AATCAA CTTGCTCTGACAGCTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCAA 5628

5221 AATCAA CTTGCTCTGACAGCTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCAA 5280

5629 ATATGAAGTGAGTGTCTATGCTCTTAAAGACACTTTTGA CAAAGCAGACCCAGCTCAGGGAGT 5688

5281 ATATGAAGTGAGTGTCTATGCTCTTAAAGACACTTTTGA CAAAGCAGACCCAGCTCAGGGAGT 5340

5689 TGTCA CCACTCTGGAGAAATGTACGCCCA CCAAGAAAGGGCTGTGTGACAGATGCTACTGA 5748

5341 TGTCA CCACTCTGGAGAAATGTACGCCCA CCAAGAAAGGGCTGTGTGACAGATGCTACTGA 5400





ACA64817;  
27-JUN-2003 (first entry)  
Human fibronectin DNA corresponding to X02761.  
Human; chronic inflammatory joint disease; infection; tumour;  
antifibronectin; cyclostatic; antirheumatic; antithrombotic;  
immunomodulatory; gene therapy; etiological pathogenicity; ds.  
Homo sapiens.  
DE10127572-A1.  
05-DEC-2002.  
30-MAY-2001; 2001DE-01027572.  
30-MAY-2001; 2001DE-01027572.  
(PATH-) PATHOARRAY GMBH.  
Haeupl T, Ungethuen U, Blaess S;  
WPI; 2003-240797/24.  
This invention describes a novel reagent for diagnosis, molecular  
definition and therapy of chronic inflammatory joint diseases, and other  
inflammatory disorders, infective or tumour diseases in humans. The  
products of the invention have antiinflammatory, cyostatic,  
antiarthritic, antirheumatic and immunosuppressive activity and can be  
used for gene therapy. The reagent of the invention and any proteins and  
antibodies derived from it, are used (i) for analysing tissue and blood  
samples for medical diagnosis; (ii) for diagnosis and characterisation of  
chronic joint diseases, on the basis of molecular characterisation, and  
determining the etiological pathogenicity principle of as yet  
uncharacterised inflammatory diseases, also monitoring progression and/or  
treatment of disease, and optimisation of therapy and (iii) for  
developing treatments for inflammatory diseases, particularly of joints,  
infections and tumours. ACA64801-ACA64965 represent human polynucleotides  
used in the method of the invention  
Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;  
Query Match 95.2%; Score 7423; DB 7; Length 7680;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
349 GAAGAGCAAGAGGAGGCTCAGCAATGTTTCAGCCCACTGCTCCCGGCTGTCAGTCA 408  
1 GAAGAGCAAGAGGAGGCTCAGCAATGTTTCAGCCCACTGCTCCCGGCTGTCAGTCA 60  
409 AAGCAAGCCCGTGTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCG 468  
61 AAGCAAGCCCGTGTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCG 120  
469 GACCTACCTAGCAATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 528  
121 GACCTACCTAGCAATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180  
529 CGAAAGTAAACCTGAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCG 588  
181 CGAAAGTAAACCTGAGCTGAAGAGACTTGTGTTGACAAAGTACACTGGGAACACTTACCG 240  
589 AGTGGGTGACACTTATGAGCGCTTAAAGACCTCATGATCTGGGACTGTACTGTCATCGG 648  
241 AGTGGGTGACACTTATGAGCGCTTAAAGACCTCATGATCTGGGACTGTACTGTCATCGG 300

QY 649 GGCTGGCGGAGGAGAAATAGCTGTACCATTCGCAAAACCGCTGCCATGAAGGGGTGAGTC 708  
DB 301 GGCTGGCGGAGGAGAAATAGCTGTACCATTCGCAAAACCGCTGCCATGAAGGGGTGAGTC 360  
QY 709 CTACAAGATTGGTGTACACCTGGAGGAGACACATGAGACTGGTGGTTACATGTTAGAGTG 768  
DB 361 CTACAAGATTGGTGTACACCTGGAGGAGACACATGAGACTGGTGGTTACATGTTAGAGTG 420  
QY 769 TGTGTGTCTTGGTAAATGGAAGAGAGAAATGAGAGCTGCAAGCCCATAGCTGAGAGTGT 828  
DB 421 TGTGTGTCTTGGTAAATGGAAGAGAGAAATGAGAGCTGCAAGCCCATAGCTGAGAGTGT 480  
QY 829 TGATCATGCTGTGGGACTTCTATGCTGCGAGAAACGTTGGAGAAACGTTGGAGAAACGTTGGAG 888  
DB 481 TGATCATGCTGTGGGACTTCTATGCTGCGAGAAACGTTGGAGAAACGTTGGAGAAACGTTGGAG 540  
QY 889 CTGGATGATGTTAGATTGTACTTCTGGGAGAAAGCGAGCGGACGATCACTTGGACTTC 948  
DB 541 CTGGATGATGTTAGATTGTACTTCTGGGAGAAAGCGAGCGGACGATCACTTGGACTTC 600  
QY 949 TAGAATGATGCAACGATCAGGACACAGGACATCTATAGAAATTGGAGACACCTGGAG 1008  
DB 601 TAGAATGATGCAACGATCAGGACACAGGACATCTATAGAAATTGGAGACACCTGGAG 660  
QY 1009 CAAGAAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 1068  
DB 661 CAAGAAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGA 720  
QY 1069 GTGGAAGTGTGAGAGGACACCTCTGTGACAGACCATCGAGCGGATCTGGCCCTTCAC 1128  
DB 721 GTGGAAGTGTGAGAGGACACCTCTGTGACAGACCATCGAGCGGATCTGGCCCTTCAC 780  
QY 1129 CGATGTTGCTGAGCTGTTTACCAACCGGAGCTCACCCCGAGCTCCCTCCCTATGGCCA 1188  
DB 781 CGATGTTGCTGAGCTGTTTACCAACCGGAGCTCACCCCGAGCTCCCTCCCTATGGCCA 840  
QY 1189 CTGTGTACAGACAGTGGTGTGTCTACTCTGTGGGGATGAGTGGTGAAGACAAAG 1248  
DB 841 CTGTGTACAGACAGTGGTGTGTCTACTCTGTGGGGATGAGTGGTGAAGACAAAG 900  
QY 1249 AAATAAGCAATCTTGTGACGCTGCTGGCAACGAGTGCAGTGCAGAGACAGCTGT 1308  
DB 901 AAATAAGCAATCTTGTGACGCTGCTGGCAACGAGTGCAGTGCAGAGACAGCTGT 960  
QY 1309 AACCCAGACTTACCGTGGCAACTCAATGAGAGCCATGTGTTTACCAATTCACCTACAA 1368  
DB 961 AACCCAGACTTACCGTGGCAACTCAATGAGAGCCATGTGTTTACCAATTCACCTACAA 1020  
QY 1369 TGGCAGGAGCTTCTACTCTCTGACACAGAGGGGCGACAGGAGCGACATCTTTGGTGCAG 1428  
DB 1021 TGGCAGGAGCTTCTACTCTCTGACACAGGAGGGGCGACAGGAGCGACATCTTTGGTGCAG 1080  
QY 1429 CACAATTCGAATTTATGAGCAGGACAGAAATCTCTTTCTGACAGACACACTGTTT 1488  
DB 1081 CACAATTCGAATTTATGAGCAGGACAGAAATCTCTTTCTGACAGACACACTGTTT 1140  
QY 1489 GGTTCAGACTCGAGGAGGAAATTCGAATGCTGCTTGTGCGACTTCCCTTCCTATACAA 1548  
DB 1141 GGTTCAGACTCGAGGAGGAAATTCGAATGCTGCTTGTGCGACTTCCCTTCCTATACAA 1200  
QY 1549 CAACCAATTTACATGATTGCACTTCTGAGGGGCGAGAGAGACAAATGAAGTGGTGTGG 1608  
DB 1201 CAACCAATTTACATGATTGCACTTCTGAGGGGCGAGAGAGACAAATGAAGTGGTGTGG 1260  
QY 1609 GACCAACAGAACTATGATGCCAGACAGAGTTGGTTCCTGCCCATGCTGCCACGA 1668  
DB 1261 GACCAACAGAACTATGATGCCAGACAGAGTTGGTTCCTGCCCATGCTGCCACGA 1320  
QY 1669 GGAATCTGCACAAACCAATGAAGGGGTCTATGACCGCAATGGAGAGTCACTGGGATGAAGCA 1728  
DB 1321 GGAATCTGCACAAACCAATGAAGGGGTCTATGACCGCAATGGAGAGTCACTGGGATGAAGCA 1380

QY 1729 GCATGACATGGGTACATGATAGGTGACGCTGTGTGGGAATGGTCGTGGGGAATGGAC 1788  
 DB 1381 GCATGACATGGGTACATGATAGGTGACGCTGTGTGGGAATGGTCGTGGGGAATGGAC 1440  
 QY 1789 ATGCAATTCCTACTCGCAGCTTCGAGATCAGTCAATTTGTTGATGACATCACTTACAAATGT 1848  
 DB 1441 ATGCAATTCCTACTCGCAGCTTCGAGATCAGTCAATTTGTTGATGACATCACTTACAAATGT 1500  
 QY 1849 GAACGACACATTCACAAAGCGTCATGAAGAGGGGCAATGCTGAACTGTACATGCTTCGG 1908  
 DB 1501 GAACGACACATTCACAAAGCGTCATGAAGAGGGGCAATGCTGAACTGTACATGCTTCGG 1560  
 QY 1909 TCAGGGTCGGGGCAGGTGGAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968  
 DB 1561 TCAGGGTCGGGGCAGGTGGAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620  
 QY 1969 GACGTTTTATCAAAATGGAGATTCATGGGAGAAATGATGTGCATGTGTGATATACCAATG 2028  
 DB 1621 GACGTTTTATCAAAATGGAGATTCATGGGAGAAATGATGTGCATGTGTGATATACCAATG 1680  
 QY 2029 CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAACCTTTACAGACCTATCCAG 2088  
 DB 1681 CTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAACCTTTACAGACCTATCCAG 1740  
 QY 2089 CTCAGGTGGTCTGTCGAAATGATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACCC 2148  
 DB 1741 CTCAGGTGGTCTGTCGAAATGATTTATCACTGAGACTCCGAGTCAGCCCAACTCCACCC 1800  
 QY 2149 CATCCAGTGAATGCACACAGCCATCTCAATTTCCAGTCAATTTCTCAGTGGAGACC 2208  
 DB 1801 CATCCAGTGAATGCACACAGCCATCTCAATTTCCAGTCAATTTCTCAGTGGAGACC 1860  
 QY 2209 TAAAAATCTGTAGGCGTTGGAGGAGGCTACATACACAGGCGCATTAACCTCCACAC 2268  
 DB 1861 TAAAAATCTGTAGGCGTTGGAGGAGGCTACATACACAGGCGCATTAACCTCCACAC 1920  
 QY 2269 CATCAAGGCGTGAAGCCTGGTGTGTATACGAGGCGAGCTCATCAGCATCCAGCAGTA 2328  
 DB 1921 CATCAAGGCGTGAAGCCTGGTGTGTATACGAGGCGAGCTCATCAGCATCCAGCAGTA 1980  
 QY 2329 CGGCCACCAAGAGTACTCGCTTGTACTTACCACACAGCAGCAGCAGCAGCTGTGAC 2388  
 DB 1981 CGGCCACCAAGAGTACTCGCTTGTACTTACCACACAGCAGCAGCAGCAGCTGTGAC 2040  
 QY 2389 CAGCAACACCGTGACAGGAGAGCAGCTCCCTTTCTCTCTTGTGGCCACTTCTGAATC 2448  
 DB 2041 CAGCAACACCGTGACAGGAGAGCAGCTCCCTTTCTCTCTTGTGGCCACTTCTGAATC 2100  
 QY 2449 TGTGACCGGAATCACAGCAGTACTGCTTGTGCTCTCGGCTCTCAGCTTCGACACCGT 2508  
 DB 2101 TGTGACCGGAATCACAGCAGTACTGCTTGTGCTCTCGGCTCTCAGCTTCGACACCGT 2160  
 QY 2509 GTCGGGATTCGGGTGGAATATGAGCTGAGTGGAGGGAGATGAGCCACAGTACCTGGA 2568  
 DB 2161 GTCGGGATTCGGGTGGAATATGAGCTGAGTGGAGGGAGATGAGCCACAGTACCTGGA 2220  
 QY 2569 TCTTCCAGACACACCACTCTGTGAACATCCCTGACCTCTCTCTGGCCGAAATACAT 2628  
 DB 2221 TCTTCCAGACACACCACTCTGTGAACATCCCTGACCTCTCTCTGGCCGAAATACAT 2280  
 QY 2629 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGATTTGATCTCTGTACTTTACA 2688  
 DB 2281 TGTAAATGTCTATCAGATATCTGAGGATGGGAGCAGATTTGATCTCTGTACTTTACA 2340  
 QY 2689 AACCAACGCGCTCATGCCCCCTCTGACCCGACCTGTGGACCAAGTGTATGACACCTCAAT 2748  
 DB 2341 AACCAACGCGCTCATGCCCCCTCTGACCCGACCTGTGGACCAAGTGTATGACACCTCAAT 2400  
 QY 2749 TGTGTTGCTGGAGCAGACCCAGGCTCCCATCACAGGCTCAGAAATAGTCTATTTCGCC 2808  
 DB 2401 TGTGTTGCTGGAGCAGACCCAGGCTCCCATCACAGGCTCAGAAATAGTCTATTTCGCC 2460  
 QY 2809 ATCAGTAGAGGTAGCAGCAGAGAACTCAACTTCTTCTGAAACTGCAAACTCCCGTCAACCT 2868

DB 2461 ATCAGTAGAGGTAGCAGCAGAGAACTCAAACTTCTTCTGAAACTGCAAACTCCGTCACCT 2520  
 QY 2869 CAGTGACTTGCACCACTGGTGTTCAGTATAAATCACTATCTATGCTGTGGAAGAAATCA 2928  
 DB 2521 CAGTGACTTGCACCACTGGTGTTCAGTATAAATCACTATCTATGCTGTGGAAGAAATCA 2580  
 QY 2929 AGAAATGACACCTGTGTTCATCAACAAAGAAACCACTGGGACCCCGCTCAGATACAGT 2988  
 DB 2581 AGAAATGACACCTGTGTTCATCAACAAAGAAACCACTGGGACCCCGCTCAGATACAGT 2640  
 QY 2989 GCCCTCTCCAGGAGACCTGCAAGTTTGTGGAGTGCACAGAGTGAAGGTCAACCATCATGTG 3048  
 DB 2641 GCCCTCTCCAGGAGACCTGCAAGTTTGTGGAGTGCACAGAGTGAAGGTCAACCATCATGTG 2700  
 QY 3049 GACACCGCTCAGAGTGCAGTGCAGGCTACCGTGTGATGTGATCCCGCTCAACCTGCC 3108  
 DB 2701 GACACCGCTCAGAGTGCAGTGCAGGCTACCGTGTGATGTGATCCCGCTCAACCTGCC 2760  
 QY 3109 TGGGAGCAGCGGAGAGGCTGCCCATCAGCAGGAAACACTTTTGCAGAGTCAACCGGCT 3168  
 DB 2761 TGGGAGCAGCGGAGAGGCTGCCCATCAGCAGGAAACACTTTTGCAGAGTCAACCGGCT 2820  
 QY 3169 GTCCCTGGGCTCACTATTACTTCAAAAGTCTTTCAGAGTGCAGTGCAGGAGGAGAGCAA 3228  
 DB 2821 GTCCCTGGGCTCACTATTACTTCAAAAGTCTTTCAGAGTGCAGTGCAGGAGGAGAGCAA 2880  
 QY 3229 GCCTCTGACTGCTCAACAGACCAACCAACTGGATGCTCCCACTAACCTCCAGTTTGTCAA 3288  
 DB 2881 GCCTCTGACTGCTCAACAGACCAACCAACTGGATGCTCCCACTAACCTCCAGTTTGTCAA 2940  
 QY 3289 TGAACATGANTTCTACTGCTGTGTGAGATGCACTCCACCTCGGCCCCAGATAAACAGATA 3348  
 DB 2941 TGAACATGANTTCTACTGCTGTGTGAGATGCACTCCACCTCGGCCCCAGATAAACAGATA 3000  
 QY 3349 CGGACTGACCTGGGCTTACTCCGAGAGGACGCGGAGGAGTCAATGTGGGTCCTC 3408  
 DB 3001 CGGACTGACCTGGGCTTACTCCGAGAGGACGCGGAGGAGTCAATGTGGGTCCTC 3060  
 QY 3409 TGTCTTCAAGTACCCACTGAGGAATCTGAGCTGCACTCTGAGTACACCTGATCCCTCGT 3468  
 DB 3061 TGTCTTCAAGTACCCACTGAGGAATCTGAGCTGCACTCTGAGTACACCTGATCCCTCGT 3120  
 QY 3469 GGCATTAAGGGGCAACCAAGAGAGCCCAAGGACCACTGTGAGTCTTTACCACTGCGACC 3528  
 DB 3121 GGCATTAAGGGGCAACCAAGAGAGCCCAAGGACCACTGTGAGTCTTTACCACTGCGACC 3180  
 QY 3529 TGGGAGCTCTATTCCACCTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATG 3588  
 DB 3181 TGGGAGCTCTATTCCACTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATG 3240  
 QY 3589 GACGCTGCTCCAGAAATGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGAGGC 3648  
 DB 3241 GACGCTGCTCCAGAAATGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGAGGC 3300  
 QY 3649 ACCAGGAGTGTCTCAGACTCAGGAAGCATGTTGTGTGCTGGCTTGTGCTCCAGAGT 3708  
 DB 3301 ACCAGGAGTGTCTCAGACTCAGGAAGCATGTTGTGTGCTGGCTTGTGCTCCAGAGT 3360  
 QY 3709 AGAATACCTCTACACCACTCCAAAGTCTGAGAGATGGAAGGAAAGAGATGCGCAATGT 3768  
 DB 3361 AGAATACCTCTACACCACTCCAAAGTCTGAGAGATGGAAGGAAAGAGATGCGCAATGT 3420  
 QY 3769 AAAAAGGTGTGACACCAATGCTCCACCAACAACTTGCATCTGGAGGCAAAACCTTGA 3828  
 DB 3421 AAAAAGGTGTGACACCAATGCTCCACCAACAACTTGCATCTGGAGGCAAAACCTTGA 3480  
 QY 3829 CACTGGAGTGTCTCAGTCTCTCTGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGT 3888  
 DB 3481 CACTGGAGTGTCTCAGTCTCTCTGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGT 3540  
 QY 3889 AATTACCAACACCTCTACAAAGCGGACGAGGAAATTTCTTTGGAAGAGTGGTCCATGC 3948



QY 6109 GGTCCCTCGCCCGCCCTGCTGTCTACAGAGGCTACTATTACTGGCTCGAACCGGAAC 6168  
DB 5761 GGTCCCTCGCCCGCCCTGCTGTCTACAGAGGCTACTATTACTGGCTCGAACCGGAAC 5820  
QY 6169 CGAATATACAAATTTATGTCTATGCTTGAAGCTGAAGATATATCAGAGAGCGAGCCCTGATTGG 6228  
DB 5821 CGAATATACAAATTTATGTCTATGCTTGAAGCTGAAGATATATCAGAGAGCGAGCCCTGATTGG 5880  
QY 6229 AAGGAAAAGACAGACAGAGCTTCCCAACCTGGTAACCCCTTCCACACCCCAATCTTCATGG 6288  
DB 5881 AAGGAAAAGACAGACAGAGCTTCCCAACCTGGTAACCCCTTCCACACCCCAATCTTCATGG 5940  
QY 6289 ACCAGAGATCTTGATGTTGCTTCCACAGTTCAAAAGACCCCTTGGTCAACCCACCCCTGG 6348  
DB 5941 ACCAGAGATCTTGATGTTGCTTCCACAGTTCAAAAGACCCCTTGGTCAACCCACCCCTGG 6000  
QY 6349 GTATGACACTGGAATGATTTAGCTTCTCGGACCTTCTGGTCAAGCAACCCAGTGTGG 6408  
DB 6001 GTATGACACTGGAATGATTTAGCTTCTCGGACCTTCTGGTCAAGCAACCCAGTGTGG 6060  
QY 6409 GCAACAAATGATCTTTGAGGAACATGTTTGGCGGACACACCCGCCCAACAGCGGCAC 6468  
DB 6061 GCAACAAATGATCTTTGAGGAACATGTTTGGCGGACACACCCGCCCAACAGCGGCAC 6120  
QY 6469 CCCCATAGGCATAGGCCAAGACCATACCCGCCAATGTAGGACAAAGACTCTCTCTCA 6528  
DB 6121 CCCCATAGGCATAGGCCAAGACCATACCCGCCAATGTAGGACAAAGACTCTCTCTCA 6180  
QY 6529 GACAAACATCTCTAGGCCCTTCCAGGACATCTTGAGTACATCATTTTCACTGATCC 6588  
DB 6181 GACAAACATCTCTAGGCCCTTCCAGGACATCTTGAGTACATCATTTTCACTGATCC 6240  
QY 6589 TGTGGCACTCATGAGAACCTTACAGTTCAGAGTTCCTGGACTTCTACCAAGTGCAC 6648  
DB 6241 TGTGGCACTCATGAGAACCTTACAGTTCAGAGTTCCTGGACTTCTACCAAGTGCAC 6300  
QY 6649 TCTGACAGGCTTACCAAGAGTGCACCTCAACATCATAGTGGAGCACTGAAAGACCA 6708  
DB 6301 TCTGACAGGCTTACCAAGAGTGCACCTCAACATCATAGTGGAGCACTGAAAGACCA 6360  
QY 6709 CGAGAGGATTAAGTTCGGAGAGGTTTACCGTGGCACTCTGTCACAGAGGCTT 6768  
DB 6361 CGAGAGGATTAAGTTCGGAGAGGTTTACCGTGGCACTCTGTCACAGAGGCTT 6420  
QY 6769 GAACCAACCTACGATGCTGCTGTGATGACCCCTTACACAGTGTCCCATTTAGCCGTGG 6828  
DB 6421 GAACCAACCTACGATGCTGCTGTGATGACCCCTTACACAGTGTCCCATTTAGCCGTGG 6480  
QY 6829 AGATGAGTGGGAACGAATGCTGAATCAGGCTTTAACTGTGTGCCAGTCTTAGGCTT 6888  
DB 6481 AGATGAGTGGGAACGAATGCTGAATCAGGCTTTAACTGTGTGCCAGTCTTAGGCTT 6540  
QY 6889 TGGAAAGTGTCTATTTTCAAGTGTGATTTCTAGATGTTGCCATGACATGTTGTGAACCTA 6948  
DB 6541 TGGAAAGTGTCTATTTTCAAGTGTGATTTCTAGATGTTGCCATGACATGTTGTGAACCTA 6600  
QY 6949 CAAGATTGAGAGAGAGTGGACCGTCAAGGAGAAATGCCAGATGATGAGCTCCACATG 7008  
DB 6601 CAAGATTGAGAGAGAGTGGACCGTCAAGGAGAAATGCCAGATGATGAGCTCCACATG 6660  
QY 7009 TCTTGGGACGCAAGAGAGAAATCAAGTGTGACCCCTCATAGGCAACGTTGTACGATGA 7068  
DB 6661 TCTTGGGACGCAAGAGAGAAATCAAGTGTGACCCCTCATAGGCAACGTTGTACGATGA 6720  
QY 7069 TGGGAGACATACCAAGTGAAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 7128  
DB 6721 TGGGAGACATACCAAGTGAAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 6780  
QY 7129 CTGCAATGCTTTGGAGGCGAGCGGCTGGCTGTGCACTGCCGACAGCTGGGGG 7188  
DB 6781 CTGCAATGCTTTGGAGGCGAGCGGCTGGCTGTGCACTGCCGACAGCTGGGGG 6840  
QY 7189 TGAACCCAGTCCCGAGGCACTACTGGCCAGTCTTACAAACCAAGTATTTCTCAGAGATACCA 7248

DB 6841 TGAACCCAGTCCCGAGGCACTACTGGCCAGTCTCTCAACCCAGTATTTCTCAGAGATACCA 6900  
QY 7249 TCAGAGAAACAAACACTAATGTTAATTTGCCCAATTTAGTGTCTTCAATGCTTTAGATGTACA 7308  
DB 6901 TCAGAGAAACAAACACTAATGTTAATTTGCCCAATTTAGTGTCTTCAATGCTTTAGATGTACA 6960  
QY 7309 GGTGACAGAGAAAGATTTCCCGAGAGTAAATCATCTTTCCCAATCCAGAGGAAACAGAGCATGT 7368  
DB 6961 GGTGACAGAGAAAGATTTCCCGAGAGTAAATCATCTTTCCCAATCCAGAGGAAACAGAGCATGT 7020  
QY 7369 CTCTCTGCAAGATCCATCTAAACTGGAGTGAATGAGTGTAGTGTAGTGTAGTGTCTTC 7428  
DB 7021 CTCTCTGCAAGATCCATCTAAACTGGAGTGAATGAGTGTAGTGTAGTGTAGTGTCTTC 7080  
QY 7429 TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCCTCAGCTTCAGCTCAACTCACAGCT 7488  
DB 7081 TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCCTCAGCTTCAGCTCAACTCACAGCT 7140  
QY 7489 TCTCCAAGCATCACCCCTGGAGTTCCTCTGAGGTTTCTCATAAATGAGGCTGTCACATT 7548  
DB 7141 TCTCCAAGCATCACCCCTGGAGTTCCTCTGAGGTTTCTCATAAATGAGGCTGTCACATT 7200  
QY 7549 GCTGTTCTGCTTCGAAGTATTTCAATACCGCTCAGTATTTTAAATGAAGTCACTTAAGA 7608  
DB 7201 GCTGTTCTGCTTCGAAGTATTTCAATACCGCTCAGTATTTTAAATGAAGTCACTTAAGA 7260  
QY 7609 TTTGGTTTGGGATCAATAGGAAGCATATGCAAGCAACCAAGATGCAAAATGTTTGAAT 7668  
DB 7261 TTTGGTTTGGGATCAATAGGAAGCATATGCAAGCAACCAAGATGCAAAATGTTTGAAT 7320  
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAGTCAACCCAAACAATCTTCTGCTTCACTTAAGTGT 7728  
DB 7321 GATATGACCAAAATTTTAAAGTAGGAAGTCAACCCAAACAATCTTCTGCTTCACTTAAGTGT 7380  
QY 7729 CTGGCCCGCAATCTGTTAGGACAGCATGATCTTGTGTACTGTGATATTTTAAATATCCA 7788  
DB 7381 CTGGCCCGCAATCTGTTAGGACAGCATGATCTTGTGTACTGTGATATTTTAAATATCCA 7440  
QY 7789 CAGTACT 7795  
DB 7441 CAGTACT 7447

Search completed: March 7, 2004, 08:59:22  
Job time : 2569 secs